IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 13, 2019, accepted January 5, 2020, date of publication January 14, 2020, date of current version January 24, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2966521

Robust Recursive Estimation for Uncertain Systems With Delayed Measurements and Noises

JIANXIN FENG [1], **RONGNI YANG** [2], **(Member, IEEE), HAIYING LIU** [1], **AND BIAO XU** [1]

[1] College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China

[2] School of Control Science and Engineering, Shandong University, Jinan 250061, China

Corresponding author: Jianxin Feng (fengjx@nuaa.edu.cn)

This work was supported by the Fundamental Research Funds for the Central Universities under Grant NS2017064.

ABSTRACT In this article, the problem of robust recursive estimation is studied for a class of uncertain systems with delayed measurements and delayed noises. The system model is subject to stochastic uncertainties which can be described by multiplicative noises. The phenomenon of delayed measurements occurs in a random way and the delay rate is characterised by a binary switch sequence with known probability distribution. The process noise and the measurement noise are both deterministic delay. By combining the noise at present time and the delayed noise into a whole one, the original system is transformed into an auxiliary stochastic uncertain system with discrete autocorrelated noises across time. Then, based on the orthogonal projection theorem and an innovation analysis approach, the desired robust recursive estimators including robust recursive filter, robust recursive predictor and robust recursive smoother are derived. A numerical simulation example is exploited to show the effectiveness of the proposed approaches.

INDEX TERMS Robust recursive estimation, delayed noise, delayed measurements, discrete autocorrelated noise, stochastic uncertainty.

I. INTRODUCTION

In the past decades, the estimation theory and design techniques have received considerable attention due to extensive application backgrounds ranging from aerospace systems, navigation, target tracking, communication systems, signal processing, and elsewhere [1]–[6]. The traditional Kalman filter is a basic and classical state-space estimator, since its inception in the early 1960s, it has attracted a great deal of attention for its simple structure and good performance. However, the good performance of the traditional Kalman filter is based on the assumption that the system structure and parameter are exactly known. Unfortunately, the uncertainty of system model is inevitable due to the complexity of the system model and the limitations of human comprehension. There are a variety of ways to describe the model uncertainties, such as the norm-bounded uncertainty [7]–[9], polytopic uncertainty [8], [10], stochastic uncertainty [9], [11]–[13] and so on. In general, the norm-bounded uncertainty and polytopic uncertainty are treated by the inequality theory and extreme value theory. For example, the iterative robust filtering problem is investigated for a certain ground target tracking system, where the norm-bounded system parameter uncertainty and input uncertainty are solved by the min-max game theory [14]. In [10], a robust $l_2 - l_\infty$ filter is designed for switched linear discrete time-delay systems with polytopic uncertainties, and the existence conditions for such a filter is formulated in terms of a set of linear matrix inequalities. The stochastic uncertainties are commonly encountered in image processing systems, communication systems and navigation systems, and are usually modeled by multiplicative noises. Different from the additive noises, the multiplicative noises are dependent on the system state, therefore, the second-order statistical properties of the multiplicative noises are usually unknown and this property leads to more difficulties in the designing of the desired estimators. Up to now, a great deal of efforts have been delivered to deal with the control and estimation problems for systems with multiplicative noises, including Riccati euqation approach [15], [16], linear matrix inequality [17] and innovation analysis approach [9], [12], [13], etc.

Most traditional estimation theories are based on the assumption that the measured data can be obtained in real time. However, this assumption is not realistic in many

The associate editor coordinating the review of this manuscript and approving it for publication was Jianyong Yao.

engineering applications. For example, in photoelectric target tracking systems, the time delay of the TV distance missing is unavoidable due to the processing of the photoelectric conversion, signal processing, data acquisition and transmission. On the other hand, with the development of the network technology, the networked control systems have been applied extensively in a lot of professions for its advantages of lows cost, great mobility and intelligence. Everything has two sides, when we enjoy the convenience brought by the network, we also have to bear the network-induced time delay. Therefore, the estimation and control problem for systems subject to delayed measurements is a hot research topic and a larger number of literature has been reported in recent years, see e.g. [11], [18]–[36]. To be specific, the problem of recursive estimation for descriptor systems with different delay rates have been investigated in [31], where the recursive filter, recursive predictor and recursive smoother are derived by using the singular value decomposition and the orthogonal projection theorem. The linear unbiased state estimation problem for one-step random sampling delay has been studied in [28], [29]. However, the estimators designed in [28], [29] are suboptimal since a colored noise due to augmentation has been treated as a white noise. Least-square linear filtering using observations coming from multiple sensors with one- or two-step random delay has also been investigated in [24], where the algorithm uses only the covariance functions of the processes involved in the observation equation of each sensor and the delay probabilities. The estimation problem for multistep time delay has also been studied in [18], [32].

It should be noted that, the past solutions to the problem of time delay mainly focus on the state time delay and measurement time delay. The cases with time delay in the noise, however, are seldom discussed. The delayed noises can be found in several research fields including biology, engineering, economics, net control systems, etc [37], [38]. In addition, the delayed noises may appear in the feedback back cases, such as, in systems with delays in the controls whenever the control action is corrupted by an additive ''white noise'' [37], [38]. The traditional method to deal with the delayed noises is through the state augmentation. However, the state augmentation will increase the system dimensions and result in expensive computational cost, especially when the time delay is large. Recently, Cui *et al.* [37] presented a new method to deal with the estimation problem with delayed noise. Different from the previous works, the new method is projection formula in Hilbert space rather than state augmentation. However, in reference [37], only the delayed process noise has been considered and the recursive predictor and recursive smoother are not designed. Up to now, to the best of the authors' knowledge, the robust recursive estimation problem has not yet been addressed for uncertain systems with delayed measurements and noises, which still remains as a challenging research issue.

Motivated by the above analysis, in this paper, we aim to investigate the robust recursive estimation problem for a class of uncertain systems with delayed measurements and noises, where the system parameters are subject to stochastic uncertainties which can be described as multiplicative noises. Without loss of generality, the measurements are assumed to be one-step time delay and the time delay phenomenon is described by a binary switching sequence that obeys a conditional probability distribution. In our current work, we do not only consider the delayed process noises, we also consider the delayed measurement noises. By combing the noise at present and the delayed noise into a whole one, the original system is transformed into an auxiliary stochastic uncertain system with discrete autocorrelated noises across time. Then, the desired robust recursive estimator including filter, predictor and smoother are obtained via the orthogonal projection theorem and an innovation analysis approach. *The main contribution of this paper is threefold: 1) the system model considered is comprehensive that takes into account the stochastic uncertainties, the randomly delayed measurements and the deterministic delayed process noises and measurement noises; 2) without resorting to state augmentation, the delayed process noises and measurement noises are treated by an innovation analysis approach and the orthogonal projection theorem; and 3) to the best of the author's knowledge, this is the first time that the discrete autocorrelated noise across time is studied in the stochastic parameter uncertain system.* A simulation example is exploited to show the effectiveness of the proposed approaches.

The remainder of the paper is organized as follows. In Section II, the problem of robust recursive estimation for a class of uncertain systems with delayed measurements and noises is formulated. The main results are derived in Section III. In Section IV, a simulation example is provided to illustrate the effectiveness of the proposed estimate schemes. We provide a conclusion in Section V.

Notation: The notation used is standard. The superscript "T" stands for matrix transposition, $\mathbb{R}^n$ denotes the n-dimensional Euclidean space, $\mathbb{R}^{m\times n}$ is the set of all real matrices of dimension $m\times n$, and I and 0 represent the identity matrix and zero matrix, respectively. The notation $diag(\ldots)$ stands for block-diagonal matrix. The notation δ_{k-j} is the Kronecker delta function, which is equal to unity for $k=j$ and zero for $k\neq j$. In addition, $\mathcal{E}\{x\}$ means mathematical expectation of x and $Prob\{\cdot\}$ represents the occurrence probability of the event "$\cdot$ ". The notation $min\{a,b\}$ means the smallest one between a and b. Matrices, if their dimensions are not explicitly stated, are assumed to be compatible for algebraic operations.

II. PROBLEM FORMULATION

Consider the following model parameter uncertain system with delayed measurements and noises:

$$\vec{x}_{k+1} = \left(\vec{A}_k + \vec{A}_{s,k}\eta_k\right)\vec{x}_k + \vec{B}_{1,k}\vec{\omega}_k + \vec{B}_{2,k-t}\vec{\omega}_{k-t}, \quad (1)$$

$$\vec{y}_k = \vec{C}_k\vec{x}_k + \vec{v}_k + \vec{v}_{k-h}, \quad (2)$$

$$y_k = (1-\lambda_k)\,\vec{y}_k + \lambda_k\vec{y}_{k-1}, \quad (3)$$

where $\vec{x}_k \in \mathbb{R}^n$ is the state of the system to be estimated, the initial value $\vec{x}_0$ has mean $\bar{\vec{x}}_0$ and covariance $\vec{P}_0 > 0$ and is uncorrelated with other noise signals. The vector $\vec{y}_k \in \mathbb{R}^m$ is the output of the sensor, and $y_k \in \mathbb{R}^m$ is the measurement received by the estimators. The vectors $\vec{\omega}_k \in \mathbb{R}^q$, $\vec{v}_k \in \mathbb{R}^m$, and $\eta_k \in \mathbb{R}^1$ are mutually uncorrelated zero-mean Gaussian white noises with covariances $\vec{Q}_k$, $\vec{R}_k$ and I respectively. The matrices $\vec{A}_k, \vec{A}_{s,k}, \vec{B}_{1,k}, \vec{B}_{2,k-t}$ and $\vec{C}_k$ are known real matrices with appropriate dimensions. The known positive integers $t > 1$ and $h > 1$ are the time delay in the process noises and measurement noises. The variable $\lambda_k \in \mathbb{R}^1$ is a binary switching sequence uncorrelated with other noise signals and has the statistic properties as follows:

$$\begin{aligned} Prob\{\lambda_k = 1\} &= \mathcal{E}\{\lambda_k\} = \beta_k, \\ Prob\{\lambda_k = 0\} &= 1 - \mathcal{E}\{\lambda_k\} = 1 - \beta_k, \end{aligned}$$

where $\beta_k \in [0 \ \ 1]$ is a known real time-varying positive scalar and λ_k is assumed to be uncorrelated with other noise signals.

By defining

$$\begin{aligned} x_k &= \begin{bmatrix} \vec{x}_k \\ \vec{x}_{k-1} \end{bmatrix}, \quad A_k = \begin{bmatrix} \vec{A}_k & 0 \\ I_n & 0 \end{bmatrix}, \ \omega_k = \begin{bmatrix} \vec{\omega}_k \\ \vec{\omega}_{k-t} \end{bmatrix}, \\ A_{s,k} &= \begin{bmatrix} \vec{A}_{s,k} & 0 \\ 0 & 0 \end{bmatrix}, \quad B_k = \begin{bmatrix} \vec{B}_{1,k} & \vec{B}_{2,k-t} \\ 0 & 0 \end{bmatrix}, \\ C_k &= [(1-\lambda_k)\vec{C}_k \ \ \lambda_k \vec{C}_{k-1}], \\ v(k) &= \begin{bmatrix} \vec{v}_k^T & \vec{v}_{k-h}^T & \vec{v}_{k-1}^T & \vec{v}_{k-1-h}^T \end{bmatrix}^T, \\ D_k &= [(1-\lambda_k)I_m \ \ (1-\lambda_k)I_m \ \ \lambda_k I_m \ \ \lambda_k I_m], \end{aligned} \tag{4}$$

a compact representation of (1)-(3) can be expressed by:

$$x_{k+1} = A_k x_k + A_{s,k}\eta_k x_k + B_k \omega_k, \tag{5}$$
$$y_k = C_k x_k + D_k v_k, \tag{6}$$

where v_k and ω_k are, respectively, the measurement noise and process noise of the system (5) -(6), and we can see from (4) that the process noise ω_k and the measurement noise v_k are zero mean and have the statistic properties $Q_{k,j}^{\omega} = \mathcal{E}\{\omega_k \omega_j^T\}$ and $R_{k,j}^{v} = \mathcal{E}\{v_k v_j^T\}$ as follows:

$$\begin{aligned} Q_{k,j}^{\omega} &= Q_k\delta_{k-j} + Q_{k,k-t}\delta_{k-t-j} + Q_{k,k+t}\delta_{k+t-j} \\ R_{k,j}^{v} &= R_k\delta_{k-j} + R_{k,k-1}\delta_{k-1-j} + R_{k,k+1}\delta_{k+1-j} \\ &\quad + R_{k,k-h}\delta_{k-h-j} + R_{k,k+h}\delta_{k+h-j} \\ &\quad + R_{k,k-1-h}\delta_{k-1-h-j} + R_{k,k+1+h}\delta_{k+1+h-j}, \end{aligned} \tag{7}$$

where

$$\begin{aligned} R_k &= diag\left(\vec{R}_k, \vec{R}_{k-h}, \vec{R}_{k-1}, \vec{R}_{k-1-h}\right), \\ Q_k &= diag(\vec{Q}_k, \vec{Q}_{k-t}), \\ Q_{k,k+t} &= \begin{bmatrix} 0 & \vec{Q}_k \\ 0 & 0 \end{bmatrix}, \quad Q_{k,k-t} = \begin{bmatrix} 0 & 0 \\ \vec{Q}_{k-t} & 0 \end{bmatrix}, \\ R_{k,k-1} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \vec{R}_{k-1} & 0 & 0 & 0 \\ 0 & \vec{R}_{k-1-h} & 0 & 0 \end{bmatrix}, \end{aligned}$$

$$\begin{aligned} R_{k,k-h} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ \vec{R}_{k-h} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \vec{R}_{k-1-h} & 0 \end{bmatrix}, \\ R_{k,k-1-h} &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \vec{R}_{k-1-h} & 0 & 0 & 0 \end{bmatrix}, \\ R_{k,k+1+h} &= \begin{bmatrix} 0 & 0 & 0 & \vec{R}_k \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ R_{k,k+1} &= \begin{bmatrix} 0 & 0 & \vec{R}_k & 0 \\ 0 & 0 & 0 & \vec{R}_{k-h} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ R_{k,k+h} &= \begin{bmatrix} 0 & \vec{R}_k & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \vec{R}_{k-1} \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned}$$

Remark 1: It can be easily seen from (7) that the process noise ω_k and the measurement noise v_k are all autocorrelated across time. For example, the process noise at time k is correlated with the process noises at time $k - t$ and $k + t$ with covariances $Q_{k,k-t}$ as well as $Q_{k,k+t}$, respectively. The measurement noise at time k is correlated with the measurement noises at time $k-1, k+1, k-h, k+h, k+h+1$ and $k - h - 1$ with covariances $R_{k,k-1}$, $R_{k,k+1}$, $R_{k,k-h}$, $R_{k,k+h}$, $R_{k,k-h-1}$ and $R_{k,k+h+1}$ respectively. Compared with the correlated noises at the same time instant, the autocorrelated noises across time will lead to more difficulties in the design of recursive estimators.

Remark 2: Furthermore, it should be pointed that the process noise ω_k and the measurement noise v_k are both discrete autocorrelated noises across time. For example, the process noise at time k is correlated with the process noise at time $k - t$, however, it is not correlated with the process noises at time $k-1, k-2, \ldots, k-t-2, k-t-1$. The measurement noise at time k is correlated with the measurement noise at time $k-1$ and $k-h$, but, it is not correlated with the measurement noise at time $k-2, k-3, \ldots, k-h-3, k-h-2$. In contrast to the continuous autocorrelated noises across time which has already been widely studied in references [2], [3], [9], [12], [13], [27], [32], the discrete autocorrelated noises across time will bring us new challenges.

Remark 3: Note that the system model of (5)-(6) is subject to stochastic uncertainties, and C_k with D_k involve the stochastic variables λ_k. Therefore, system (5)-(6) is actually a stochastic system. On the other hand, the measurement noise v_k and the process noise ω_k are both discrete autocorrelated noise across. In view of these two observations, conventional robust recursive estimators are no longer applicable here.

The state estimation is a dynamic estimation problem, which can be divided into three types. The filtering problem is to use the current measurement information to estimate the current system state, the prediction problem is to use the current measurement information to estimate the future system state, and the smoother problem is to use the current measurement information to estimate the past system state. Our aim in this paper is to design recursive filter $\hat{x}_{k|k}$, predictor $\hat{x}_{k+N|k}, N > 1$ and smoother $\hat{x}_{k|k+N}, N \geq 1$ for the given system. For this purpose, the orthogonal projection theorem and an innovation analysis approach will be used. The advantage of this proposed method to address the recursive estimation comes from the fact that the innovations constitute a white process and the newly obtained estimators are optimal in the linear minimum variance sense.

III. MAIN RESULTS

Before processing further, let us introduce some new notations and Lemmas, which are very useful in establishing our main results.

$$\begin{aligned} \bar{C}_{e,k} &= [-\vec{C}_k \;\; \vec{C}_{k-1}], \quad J_k = (\lambda_k - \beta_k), \\ \bar{D}_e &= [-I_m \;\; -I_m \;\; I_m \;\; I_m], \quad \sigma_k = \mathcal{E}\left\{J_k^2\right\} \\ \tilde{C}_k &= C_k - \bar{C}_k = J_k \bar{C}_{e,k}, \quad \tilde{D}_k = J_k \bar{D}_e. \end{aligned} \tag{8}$$

Lemma 1: For system state x_k and process noise $\omega_{k+m}(m \geq 0)$, the second order mixed origin moment $X^{x,\omega}_{k,k+m} = \mathcal{E}\{x_k \omega^T_{k+m}\}$ can be calculated as follows:

$$X^{x,\omega}_{k,k+m} = \begin{cases} \digamma_{t-m-1} B_{k+m-t} \times Q^T_{k+m,k+m-t}, & 0 \leq m < t, \\ 0, & m \geq t, \end{cases}$$

where the notation $\digamma_i$ is defined as follows:

$$\digamma_i = \begin{cases} \prod_{f=1}^{i} A_{k-f}, & i > 0, \\ I, & i = 0. \end{cases} \tag{9}$$

Proof: From (5) and (9), the system state x_k can be rewritten as follows:

$$x_k = \prod_{i=1}^{t} A_{k-i} x_{k-t} + \sum_{i=1}^{t} \digamma_{i-1} \times (A_{s,k-i}\eta_{k-i}x_{k-i} + B_{k-i}\omega_{k-i}). \tag{10}$$

Taking into account (10) and the fact that η_{k-i} is not correlated with ω_{k+m}, the mixed origin moment $X^{x,\omega}_{k,k+m}$ can be calculated as follows:

$$X^{x,\omega}_{k,k+m} = \prod_{i=1}^{t} A_{k-i}\mathcal{E}\{x_{k-t}\omega^T_{k+m}\} + \sum_{i=1}^{t} \digamma_{i-1} B_{k-i}\mathcal{E}\{\omega_{k-i}\omega^T_{k+m}\}.$$

It follows readily from (5), (7) and $k+m-k+t = m+t \geq t$ that the system state x_{k-t} is not correlated with the process noise ω_{k+m}. Therefore, we have

$$X^{x,\omega}_{k,k+m} = \sum_{i=1}^{t} \digamma_{i-1} B_{k-i}\mathcal{E}\{\omega_{k-i}\omega^T_{k+m}\}. \tag{11}$$

If we want $\mathcal{E}\{\omega_{k-i}\omega^T_{k+m}\} \neq 0$, the subscripts of ω_{k-i} and ω_{k+m} should meet the following relationship:

$$k + m - k + i = m + i = t, \quad i \in \{1, 2, \ldots, t\}. \tag{12}$$

If $m \geq t$, then (12) does not hold, that is to say the expectation $\mathcal{E}\{\omega_{k-i}\omega^T_{k+m}\} = 0$, further more, the mixed origin moment $X^{x,\omega}_{k,k+m} = 0$. If $0 \leq m < t$, then (12) holds, and (12) can be rewritten as follows:

$$i = t - m, \tag{13}$$

where the value of i is set {1,2,...,t}, however, the values of m and t are fixed and unique. Therefore, the value of i which satisfies (12) and (13) is fixed and unique. Substituting (13) into (11), we have

$$X^{x,\omega}_{k,k+m} = \digamma_{t-m-1} B_{k+m-t} Q^T_{k+m,k+m-t}, \quad 0 \leq m < t, \tag{14}$$

which completes the proof of Lemma 1.

Remark 4: If the process noise ω_k is continuous autocorrelated across time and if we want the expectation $\mathcal{E}\{\omega_{k-i}\omega^T_{k+m}\} \neq 0$, then, equation (12) will be changed as $m + i \leq t$ and the variable i in (13) are not unique. Therefore, equations (12)-(14) constitute the main differences between the discrete autocorrelated process noise across time and the continuous autocorrelated process noise across time in the proof of the second order mixed origin moment $X^{x,\omega}_{k,k+m}$.

Lemma 2: For system state second origin moment matrix $X^{x,x}_{k+1,k+1} = \mathcal{E}\{x_{k+1}x^T_{k+1}\}$, we have the following recursive result:

$$X^{x,x}_{k+1,k+1} = A_k X^{x,x}_{k,k} A^T_k + A_k X^{x,\omega}_{k,k} B^T_k + A_{s,k} X^{x,x}_{k,k} \times A^T_{s,k} + B_k (X^{x,\omega}_{k,k})^T A^T_k + B_k Q_k B^T_k,$$

where $X^{x,\omega}_{k,k}$ can be calculated as in Lemma 1 and the initial value is $X^{x,x}_{0,0} = diag(\bar{\vec{x}}_0 \bar{\vec{x}}^T_0, 0) + diag(\vec{P}_0, 0)$.

Proof: Lemma 2 follows directly from (5), Lemma 1 and the fact that the noise signal η_k is zero mean unit variance and uncorrelated with other signals.

Lemma 3: The innovation ε_k, the process noise one-step predictor $\hat{\omega}_{k|k-1}$ and the measurement noise on-step predictor $\hat{v}_{k|k-1}$ are given by:

$$\varepsilon_k = y_k - \bar{C}_k \hat{x}_{k|k-1} - \bar{D}_k \hat{v}_{k|k-1}, \tag{15}$$

$$\hat{\omega}_{k|k-1} = \sum_{i=1}^{t-1} \Xi^{\omega,\varepsilon}_{k,k-i} \Lambda^{-1}_{k-i} \varepsilon_{k-i}, \tag{16}$$

$$\hat{v}_{k|k-1} = \sum_{i=1}^{h+1} \Xi^{v,\varepsilon}_{k,k-i} \Lambda^{-1}_{k-i} \varepsilon_{k-i}, \tag{17}$$

where $\hat{x}_{k|k-1}$ is the system state one-step predictor, the innovation covariance Λ_{k-i} will be determined as in Theorem 1, the expectations $\Xi^{\omega,\varepsilon}_{k,k-i} = \mathcal{E}\{\omega_k \varepsilon^T_{k-i}\}$ and $\Xi^{v,\varepsilon}_{k,k-i} = \mathcal{E}\{v_k \varepsilon^T_{k-i}\}$ can be calculated as follows:

$$\Xi^{\omega,\varepsilon}_{k,k-i} = \Theta^{\omega,x}_{k,k-i|k-i-1} \bar{C}^T_{k-i} + \Theta^{\omega,v}_{k,k-i|k-i-1} \bar{D}^T_{k-i},$$

$$\Xi^{v,\varepsilon}_{k,k-i} = \Theta^{v,x}_{k,k-i|k-i-1}\bar{C}^T_{k-i} + \Theta^{v,v}_{k,k-i|k-i-1}\bar{D}^T_{k-i},$$

where the expectations $\Theta^{\omega,x}_{k,k-i|k-i-1} = \mathcal{E}\left\{\omega_k\tilde{x}^T_{k-i|k-i-1}\right\}$, $\Theta^{\omega,v}_{k,k-i|k-i-1} = \mathcal{E}\left\{\omega_k\tilde{v}^T_{k-i|k-i-1}\right\}$, $\Theta^{v,x}_{k,k-i|k-i-1} = \mathcal{E}\left\{v_k\tilde{x}^T_{k-i|k-i-1}\right\}$ and $\Theta^{v,v}_{k,k-i|k-i-1} = \mathcal{E}\left\{v_k\tilde{v}^T_{k-i|k-i-1}\right\}$ are, respectively, determined by:

$$\begin{aligned}\Theta^{\omega,x}_{k,k-i|k-i-1} &= \prod_{n=i+1}^{t-1} Q_{k,k-t}B^T_{k-t}\mathcal{A}^T_{k-n} \\ &\quad + \sum_{n=i+1}^{t-1}\Big(\Theta^{\omega,\omega}_{k,k-n|k-n-1}B^T_{k-n} \\ &\quad + \Theta^{\omega,v}_{k,k-n|k-n-1}\mathcal{D}^T_{k-n}\Big)\mathbb{A}^T_n, \qquad (18)\\ \Theta^{\omega,v}_{k,k-i|k-i-1} &= -\sum_{f=i+1}^{M_1}\Xi^{\omega,\varepsilon}_{k,k-f}\Lambda^{-1}_{k-f}(\Xi^{v,\varepsilon}_{k-i,k-f})^T,\\ M_1 &= min\{h+i+1, t-1\}, \qquad (19)\\ \Theta^{v,x}_{k,k-i|k-i-1} &= \prod_{n=i+1}^{h}\Big(-R_{k,k-h-1}\bar{D}^T_{k-h-1} \\ &\quad \times \Lambda^{-1}_{k-h-1}(\Xi^{x,\varepsilon}_{k-h,k-h-1})^T\Big)\mathcal{A}^T_{k-n} \\ &\quad + \sum_{n=i+1}^{h}\Big(\Theta^{v,\omega}_{k,k-n|k-n-1}B^T_{k-n} \\ &\quad + \Theta^{v,v}_{k,k-n|k-n-1}\mathcal{D}^T_{k-n}\Big)\mathbb{A}^T_n, \qquad (20)\\ \Theta^{v,v}_{k,k-i|k-i-1} &= R^v_{k,k-i} \\ &\quad - \sum_{f=i+1}^{h+1}\Xi^{v,\varepsilon}_{k,k-f}\Lambda^{-1}_{k-f}(\Xi^{v,\varepsilon}_{k-i,k-f})^T, \qquad (21)\end{aligned}$$

where the matrices $\mathcal{A}^T_{k-n}$, $\mathcal{D}^T_{k-n}$ and $\mathbb{A}^T_n$ are defined in (48) and (28), respectively. The remaining expectations $\Theta^{\omega,\omega}_{k,k-n|k-n-1} = \mathcal{E}\{\omega_k\tilde{\omega}^T_{k-n|k-n-1}\}$ and $\Theta^{v,\omega}_{k,k-n|k-n-1} = \mathcal{E}\{v_k\tilde{\omega}^T_{k-n|k-n-1}\}$ can be calculated as follows:

$$\begin{aligned}\Theta^{\omega,\omega}_{k,k-n|k-n-1} &= Q^{\omega}_{k,k-n} - \sum_{f=n+1}^{t+n-1}\Xi^{\omega,\varepsilon}_{k,k-f} \\ &\quad \times \Lambda^{-1}_{k-f}(\Xi^{\omega,\varepsilon}_{k-n,k-f})^T, \qquad (22)\\ \Theta^{v,\omega}_{k,k-n|k-n-1} &= -\sum_{f=n+1}^{M_2}\Xi^{v,\varepsilon}_{k,k-f}\Lambda^{-1}_{k-f}(\Xi^{\omega,\varepsilon}_{k-n,k-f})^T,\\ M_2 &= min\{h+1, t+n-1\}, \qquad (23)\end{aligned}$$

where the initial values are $\varepsilon_{-i} = 0$, $\Xi^{v,\varepsilon}_{0,-i} = 0$, $\Xi^{\omega,\varepsilon}_{0,-i} = 0$, $\Xi^{v,\varepsilon}_{-n,-i} = 0$, $\Xi^{\omega,\varepsilon}_{-n,-i} = 0$, $\Lambda_{-i} = I_m$, $\Theta^{\omega,\omega}_{0,0-n|0-n-1} = 0$, $\Theta^{v,\omega}_{0,0-i|0-i-1} = 0$, $\Theta^{v,v}_{0,0-i|0-i-1} = 0$, $\Theta^{v,x}_{0,0-i|0-i-1} = 0$, $\Theta^{x,\omega}_{0,0-i|0-i-1} = 0$ and $\hat{x}_{0|-1} = [\bar{x}^T_0 \quad 0]^T$, and the range of the values of i and n are defined as in (18)-(23).

Proof: It follows from (6), (7) and the OPT that the one-step prediction for y_k can be calculated as follows:

$$\begin{aligned}\hat{y}_{k|k-1} &= \bar{C}_k\mathcal{E}\{x_k\} + \bar{C}_k\sum_{i=1}^{k-1}\mathcal{E}\{x_k\varepsilon^T_i\}\Lambda^{-1}_i\varepsilon_i \\ &\quad + \bar{D}_k\sum_{i=1}^{k-1}\mathcal{E}\{v_k\varepsilon^T_i\}\Lambda^{-1}_i\varepsilon_i \\ &= \bar{C}_k\hat{x}_{k|k-1} + \bar{D}_k\hat{v}_{k|k-1}. \qquad (24)\end{aligned}$$

Subtracting (24) from y_k yields (15).

Applying the OPT, the process noise one-step predictor $\hat{\omega}_{k|k-1}$ and the measurement noise one-step predictor $\hat{v}_{k|k-1}$ can be calculated as follows:

$$\begin{aligned}\hat{\omega}_{k|k-1} &= \sum_{i=1}^{k-1}\mathcal{E}\{\omega_k\varepsilon^T_{k-i}\}\Lambda^{-1}_{k-i}\varepsilon_{k-i},\\ \hat{v}_{k|k-1} &= \sum_{i=1}^{k-1}\mathcal{E}\{v_k\varepsilon^T_{k-i}\}\Lambda^{-1}_{k-i}\varepsilon_{k-i}.\end{aligned}$$

From (7), we know that when $i > t-1$, the expectation $\mathcal{E}\{\omega_k\varepsilon^T_{k-i}\} = 0$, and when $i > h+1$, the expectation $\mathcal{E}\{v_k\varepsilon^T_{k-i}\} = 0$. Thus, the process noise one-step predictor $\hat{\omega}_{k|k-1}$ and the measurement noise one-step predictor $\hat{v}_{k|k-1}$ can be rewritten as in (16) and (17), respectively.

Taking into account (17) and the fact that the process noise ω_k is not correlated with the measurement noise v_k, the expectation $\Theta^{\omega,v}_{k,k-i|k-i-1} = \mathcal{E}\{\omega_k\tilde{v}^T_{k-i|k-i-1}\}$ can be calculated as follows:

$$\begin{aligned}\Theta^{\omega,v}_{k,k-i|k-i-1} &= \mathcal{E}\left\{\omega_k v^T_{k-i}\right\} - \mathcal{E}\left\{\omega_k\hat{v}^T_{k-i|k-i-1}\right\} \\ &= -\sum_{f=1+i}^{h+i+1}\mathcal{E}\{\omega_k\varepsilon^T_{k-f}\}\Lambda^{-1}_{k-f} \\ &\quad \times \mathcal{E}\{v_{k-i}\varepsilon^T_{k-f}\}^T. \qquad (25)\end{aligned}$$

If $f > t-1$, then the expectation $\mathcal{E}\{\omega_k\varepsilon^T_{k-f}\} = 0$ and we have (19). Similarly, $\Theta^{v,v}_{k,k-i|k-i-1} = \mathcal{E}\{v_k\tilde{v}^T_{k-i|k-i-1}\}$ can be calculated as follows:

$$\begin{aligned}\Theta^{v,v}_{k,k-i|k-i-1} &= \mathcal{E}\left\{v_k v^T_{k-i}\right\} - \mathcal{E}\left\{v_k\hat{v}^T_{k-i|k-i-1}\right\} \\ &= R^v_{k,k-i} - \sum_{f=1+i}^{h+1}\Xi^{v,\varepsilon}_{k,k-f} \\ &\quad \times \Lambda^{-1}_{k-f}(\Xi^{v,\varepsilon}_{k-i,k-f})^T. \qquad (26)\end{aligned}$$

According to (6), (8) and (15), the innovation ε_k can be rewritten as follows:

$$\varepsilon_k = \tilde{C}_k x_k + \bar{C}_k\tilde{x}_{k|k-1} + \tilde{D}_k v_k + \bar{D}_k\tilde{v}_{k|k-1}. \qquad (27)$$

Therefore, using (25), (26) and the fact that $\tilde{D}_{k-i}$ and $\tilde{C}_{k-i}$ are both zero mean and uncorrelated with other signals, the expectations $\Xi^{\omega,\varepsilon}_{k,k-i} = \mathcal{E}\{\omega_k \varepsilon^T_{k-i}\}$ and $\Xi^{v,\varepsilon}_{k,k-i} = \mathcal{E}\{v_k \varepsilon^T_{k-i}\}$ can be calculated as follows:

$$\begin{aligned}
\Xi^{\omega,\varepsilon}_{k,k-i} &= \mathcal{E}\left\{\omega_k \left(\tilde{C}_{k-i}x_{k-i} + \bar{C}_{k-i}\tilde{x}_{k-i|k-i-1} + \tilde{D}_{k-i}v_{k-i} + \bar{D}_{k-i}\tilde{v}_{k-i|k-i-1}\right)^T\right\}\\
&= \mathcal{E}\left\{\omega_k \tilde{x}^T_{k-i|k-i-1}\right\}\bar{C}^T_{k-i} + \mathcal{E}\left\{\omega_k \tilde{v}^T_{k-i|k-i-1}\right\}\bar{D}^T_{k-i}\\
&= \Theta^{\omega,x}_{k,k-i|k-i-1}\bar{C}^T_{k-i} + \Theta^{\omega,v}_{k,k-i|k-i-1}\bar{D}^T_{k-i},\\
\Xi^{v,\varepsilon}_{k,k-i} &= \mathcal{E}\left\{v_k \left(\tilde{C}_{k-i}x_{k-i} + \bar{C}_{k-i}\tilde{x}_{k-i|k-i-1} + \tilde{D}_{k-i}v_{k-i} + \bar{D}_{k-i}\tilde{v}_{k-i|k-i-1}\right)^T\right\}\\
&= \mathcal{E}\left\{v_k \tilde{x}^T_{k-i|k-i-1}\right\}\bar{C}^T_{k-i} + \mathcal{E}\left\{v_k \tilde{v}^T_{k-i|k-i-1}\right\}\bar{D}^T_{k-i}\\
&= \Theta^{v,x}_{k,k-i|k-i-1}\bar{C}^T_{k-i} + \Theta^{v,v}_{k,k-i|k-i-1}\bar{D}^T_{k-i}.
\end{aligned}$$

Then, our next step is to calculate the remaining expectations $\Theta^{\omega,x}_{k,k-i|k-i-1}$ and $\Theta^{v,x}_{k,k-i|k-i-1}$.

From Theorem 1, the state prediction error $\tilde{x}_{k-i|k-i-1}$ can be calculated as follows:

$$\begin{aligned}
\tilde{x}_{k-i|k-i-1} =\ & \mathcal{A}_{k-i-1}\tilde{x}_{k-i-1|k-i-2} + \mathcal{A}_{s,k-i-1}\\
&+ B_{k-i-1}\tilde{\omega}_{k-i-1|k-i-2}\\
&- \mathcal{D}_{k-i-1}\tilde{v}_{k-i-1|k-i-2},
\end{aligned}$$

where $\mathcal{A}_{k-i-1}$, $\mathcal{A}_{s,k-i-1}$ and $\mathcal{D}_{k-i-1}$ are determined as in (48). By introducing the notation $\mathbb{A}_n$

$$\mathbb{A}_n = \begin{cases} \prod_{j=i+2}^{n}\mathcal{A}_{k-j+1}, & 0 \le j \le n,\\ I, & j > n, \end{cases} \tag{28}$$

the state prediction error $\tilde{x}_{k-i|k-i-1}$ can be rewritten by:

$$\begin{aligned}
\tilde{x}_{k-i|k-i-1} =\ & \prod_{n=i+1}^{t-1}\mathcal{A}_{k-n}\tilde{x}_{k-t+1|k-t}\\
&+ \sum_{n=i+1}^{t-1}\mathbb{A}_n\left\{\mathcal{A}_{s,k-n} + B_{k-n}\tilde{\omega}_{k-n|k-n-1} - \mathcal{D}_{k-n}\tilde{v}_{k-n|k-n-1}\right\},
\end{aligned}$$

therefore, the expectations $\Theta^{\omega,x}_{k,k-i|k-i-1} = \mathcal{E}\left\{\omega_k \tilde{x}^T_{k-i|k-i-1}\right\}$ and $\Theta^{v,x}_{k,k-i|k-i-1} = \mathcal{E}\left\{v_k \tilde{x}^T_{k-i|k-i-1}\right\}$ can be determined as follows:

$$\begin{aligned}
\Theta^{\omega,x}_{k,k-i|k-i-1} =\ & \prod_{n=i+1}^{t-1}\mathcal{E}\left\{\omega_k \tilde{x}^T_{k-t+1|k-t}\right\}\mathcal{A}^T_{k-n}\\
&+ \sum_{n=i+1}^{t-1}\Big(\mathcal{E}\{\omega_k \mathcal{A}^T_{s,k-n}\}\\
&+ \mathcal{E}\{\omega_k \tilde{\omega}^T_{k-n|k-n-1}B^T_{k-n}\}\\
&- \mathcal{E}\{\omega_k \tilde{v}^T_{k-n|k-n-1}\mathcal{D}^T_{k-n}\}\Big)\mathbb{A}^T_n,
\end{aligned} \tag{29}$$

$$\begin{aligned}
\Theta^{v,x}_{k,k-i|k-i-1} =\ & \prod_{n=i+1}^{h}\mathcal{E}\left\{v_k \tilde{x}^T_{k-h|k-h-1}\right\}\mathcal{A}^T_{k-n}\\
&+ \sum_{n=i+1}^{h}\Big(\mathcal{E}\{v_k \mathcal{A}^T_{s,k-n}\}\\
&+ \mathcal{E}\{v_k \tilde{\omega}^T_{k-n|k-n-1}B^T_{k-n}\}\\
&+ \mathcal{E}\{v_k \tilde{v}^T_{k-n|k-n-1}\mathcal{D}^T_{k-n}\}\Big)\mathbb{A}^T_n.
\end{aligned} \tag{30}$$

Taking into account (48), Lemma 1 and the fact that ω_k and v_k are both discrete autocorrelated across time, the remaining expectations in (29) and (30) can be calculated as follows:

$$\begin{aligned}
\mathcal{E}\left\{\omega_k \tilde{x}^T_{k-t+1|k-t}\right\} &= \mathcal{E}\left\{\omega_k (x_{k-t+1} - \hat{x}_{k-t+1|k-t})^T\right\}\\
&= Q_{k,k-t}B^T_{k-t}, &(31)\\
\mathcal{E}\{\omega_k \mathcal{A}^T_{s,k-n}\} &= 0, &(32)\\
\Theta^{\omega,\omega}_{k,k-n|k-n-1} &= \mathcal{E}\{\omega_k \omega^T_{k-n}\} - \mathcal{E}\{\omega_k \hat{\omega}^T_{k-n|k-n-1}\}\\
&= Q^{\omega}_{k,k-n} - \sum_{f=n+1}^{t-1}\Xi^{\omega,\varepsilon}_{k,k-f}\\
&\quad\times \Lambda^{-1}_{k-f}(\Xi^{\omega,\varepsilon}_{k-n,k-f})^T, &(33)\\
\mathcal{E}\left\{v_k \tilde{x}^T_{k-h|k-h-1}\right\} &= \mathcal{E}\left\{v_k x^T_{k-h}\right\} - \mathcal{E}\left\{v_k \hat{x}^T_{k-h|k-h-1}\right\}\\
&= -R_{k,k-h-1}\bar{D}^T_{k-h-1}\\
&\quad\times \Lambda^{-1}_{k-h-1}(\Xi^{x,\varepsilon}_{k-h,k-h-1})^T, &(34)\\
\mathcal{E}\{v_k \mathcal{A}^T_{s,k-n}\} &= 0, &(35)\\
\Theta^{v,\omega}_{k,k-n|k-n-1} &= \mathcal{E}\{v_k \hat{\omega}^T_{k-n|k-n-1}\} + \mathcal{E}\{v_k \omega^T_{k-n}\}\\
&= -\sum_{f=n+1}^{M_2}\Xi^{v,\varepsilon}_{k,k-f}\Lambda^{-1}_{k-f}(\Xi^{\omega,\varepsilon}_{k-n,k-f})^T,\\
M_2 &= min\{h+1, t+n-1\}. &(36)
\end{aligned}$$

Combining (31)-(33) and (29), we have (18). Substituting (34)-(36) into (30) yields (20) which completes the proof.

Remark 5: In the traditional recursive estimation problem, the innovations are calculated as $\varepsilon_k = y_k - C_k\hat{x}_{k|k-1}$ and the noises one-step predictors are $\hat{\omega}_{k|k-1} = 0$ and $\hat{v}_{k|k-1} = 0$. However, due to the possible sensor delay which occurs in a random way and the deterministic delayed noises, these are not true for the problem at hand. Therefore, we need to recalculate the innovations, the process noise one-step predictor and the measurement noise one-step predictor.

Lemma 4: For the state one-step prediction errors $\tilde{r}_{k|k-1}$ and $\tilde{z}_{k|k-1}$, we have the following result:

$$\mathcal{E}\{\tilde{r}_{k|k-1}\tilde{z}^T_{k|k-1}\} = \mathcal{E}\{r_k \tilde{z}^T_{k|k-1}\}.$$

Proof: According to the OPT, the one-step predictors $\hat{r}_{k|k-1}$ and $\hat{z}_{k|k-1}$ can be calculated as follows:

$$\begin{aligned}\hat{r}_{k|k-1} &= \mathcal{E}\{r_k\} + \sum_{i=1}^{k-1}\mathcal{E}\{r_k\varepsilon_i^T\}\Lambda_i^{-1}\varepsilon_i \\ \hat{z}_{k|k-1} &= \mathcal{E}\{z_k\} + \sum_{j=1}^{k-1}\mathcal{E}\{z_k\varepsilon_j^T\}\Lambda_j^{-1}\varepsilon_j,\end{aligned}$$

therefore, the one-step prediction error $\tilde{r}_{k|k-1} = r_k - \hat{r}_{k|k-1}$ and $\tilde{z}_{k|k-1} = z_k - \hat{z}_{k|k-1}$ can be calculated as follows:

$$\begin{aligned}\tilde{r}_{k|k-1} &= r_k - \mathcal{E}\{r_k\} - \sum_{i=1}^{k-1}\mathcal{E}\{r_k\varepsilon_i^T\}\Lambda_i^{-1}\varepsilon_i, \\ \tilde{z}_{k|k-1} &= z_k - \mathcal{E}\{z_k\} - \sum_{j=1}^{k-1}\mathcal{E}\{z_k\varepsilon_j^T\}\Lambda_j^{-1}\varepsilon_j. \end{aligned} \tag{37}$$

It follows directly from (37) that the expectation $\mathcal{E}\{\tilde{r}_{k|k-1}\tilde{z}_{k|k-1}^T\}$ can be calculated as follows:

$$\begin{aligned}\mathcal{E}\{\tilde{r}_{k|k-1}\tilde{z}_{k|k-1}^T\} &= \mathcal{E}\{r_k\tilde{z}_{k|k-1}^T\} - \mathcal{E}\{\mathcal{E}\{r_k\}z_k^T\} \\ &\quad + \mathcal{E}\{\mathcal{E}\{r_k\}\mathcal{E}\{z_k\}^T\} \\ &\quad + \sum_{j=1}^{k-1}\mathcal{E}\{\mathcal{E}\{r_k\}\varepsilon_j^T\}\Lambda_j^{-1}\mathcal{E}\{z_k\varepsilon_j^T\}^T \\ &\quad - \sum_{i=1}^{k-1}\mathcal{E}\{r_k\varepsilon_i^T\}\Lambda_i^{-1}\mathcal{E}\{\varepsilon_i z_k^T\} \\ &\quad + \sum_{i=1}^{k-1}\mathcal{E}\{r_k\varepsilon_i^T\}\Lambda_i^{-1}\mathcal{E}\{\varepsilon_i\mathcal{E}\{z_k\}^T\} \\ &\quad + \sum_{j=1}^{k-1}\sum_{i=1}^{k-1}\mathcal{E}\{r_k\varepsilon_i^T\}\Lambda_i^{-1}\mathcal{E}\{\varepsilon_i\varepsilon_j^T\} \\ &\quad \times \Lambda_j^{-1}\mathcal{E}\{\varepsilon_j z_k^T\}, \end{aligned} \tag{38}$$

since the innovations are zero mean and uncorrelated with each other, in addition, the expectations $\mathcal{E}\{r_k\}$ and $\mathcal{E}\{z_k\}$ are uncorrelated with the innovations ε_i, therefore, we have

$$\begin{aligned}\mathcal{E}\{\mathcal{E}\{r_k\}z_k^T\} &= \mathcal{E}\{\mathcal{E}\{r_k\}\mathcal{E}\{z_k\}^T\} = \mathcal{E}\{r_k\}\mathcal{E}\{z_k\}^T, \\ \mathcal{E}\{\mathcal{E}\{r_k\}\varepsilon_j^T\} &= \mathcal{E}\{r_k\}\mathcal{E}\{\varepsilon_j^T\} = 0, \\ \mathcal{E}\{\varepsilon_i\mathcal{E}\{z_k\}^T\} &= \mathcal{E}\{\varepsilon_i\}\mathcal{E}\{z_k^T\} = 0, \\ \sum_{j=1}^{k-1}\sum_{i=1}^{k-1}\mathcal{E}\{r_k\varepsilon_i^T\}&\Lambda_i^{-1}\mathcal{E}\{\varepsilon_i\varepsilon_j^T\}\Lambda_j^{-1}\mathcal{E}\{\varepsilon_j z_k^T\} \\ &= \sum_{i=1}^{k-1}\mathcal{E}\{r_k\varepsilon_i^T\}\Lambda_i^{-1}\mathcal{E}\{\varepsilon_i z_k^T\}. \end{aligned} \tag{39}$$

Substituting (39) into (38), we have $\mathcal{E}\{\tilde{r}_{k|k-1}\tilde{z}_{k|k-1}^T\} = \mathcal{E}\{r_k\tilde{z}_{k|k-1}^T\}$ which completes the proof.

A. ROBUST RECURSIVE FILTER

Theorem 1: For system (5)-(6), we have the following robust recursive filter:

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \Xi_{k,k}^{x,\varepsilon}\Lambda_k^{-1}\varepsilon_k, \tag{40}$$

$$P_{k|k} = P_{k|k-1} - \Xi_{k,k}^{x,\varepsilon}\Lambda_k^{-1}(\Xi_{k,k}^{x,\varepsilon})^T, \tag{41}$$

$$\hat{x}_{k+1|k} = \hat{x}_{k+1|k-1} + \Xi_{k+1,k}^{x,\varepsilon}\Lambda_k^{-1}\varepsilon_k, \tag{42}$$

$$\begin{aligned}P_{k+1|k} &= \mathcal{A}_k P_{k|k-1}\mathcal{A}_k^T + \mathcal{A}_k(\Theta_{k,k|k-1}^{\omega,x})^T B_k^T \\ &\quad - \mathcal{A}_k(\Theta_{k,k|k-1}^{v,x})^T\mathcal{D}_k^T + X_{k,k}^{\mathcal{A}_s,\mathcal{A}_s} \\ &\quad + B_k\Theta_{k,k|k-1}^{\omega,x}\mathcal{A}_k^T + B_k\Theta_{k,k|k-1}^{\omega,\omega}B_k^T \\ &\quad - B_k(\Theta_{k,k|k-1}^{v,\omega})^T\mathcal{D}_k^T - \mathcal{D}_k\Theta_{k,k|k-1}^{v,x}\mathcal{A}_k^T \\ &\quad - \mathcal{D}_k\Theta_{k,k|k-1}^{v,\omega}B_k^T + \mathcal{D}_k\Theta_{k,k|k-1}^{v,v}\mathcal{D}_k^T, \end{aligned} \tag{43}$$

$$\begin{aligned}\Lambda_k &= \sigma_k\bar{C}_{e,k}X_{k,k}^{x,x}\bar{C}_{e,k}^T + \bar{C}_k P_{k|k-1}\bar{C}_k^T \\ &\quad + \sigma_k\bar{D}_e R_k\bar{D}_e^T + \bar{D}_k\Theta_{k,k|k-1}^{v,v}\bar{D}_k^T, \end{aligned} \tag{44}$$

$$\Xi_{k,k}^{x,\varepsilon} = P_{k|k-1}\bar{C}_k^T + (\Theta_{k,k|k-1}^{v,x})^T\bar{D}_k^T, \tag{45}$$

$$\Xi_{k,k+1}^{x,\varepsilon} = A_k\Xi_{k,k}^{x,\varepsilon} + B_k\Xi_{k,k}^{\omega,\varepsilon}, \tag{46}$$

$$\begin{aligned}X_{k,k}^{\mathcal{A}_s,\mathcal{A}_s} &= A_{s,k}X_{k,k}^{x,x}A_{s,k}^T + \sigma_k\Xi_{k+1,k}^{x,\varepsilon}\Lambda_k^{-1}\bar{D}_e R_k \\ &\quad \times \bar{D}_e^T\Lambda_k^{-1}(\Xi_{k+1,k}^{x,\varepsilon})^T + \sigma_k\Xi_{k+1,k}^{x,\varepsilon} \\ &\quad \times \Lambda_k^{-1}\bar{C}_{e,k}X_{k,k}^{x,x}\bar{C}_{e,k}^T\Lambda_k^{-1}(\Xi_{k+1,k}^{x,\varepsilon})^T, \end{aligned} \tag{47}$$

where the innovation ε_k is defined and calculated as in Lemma 3. Λ_k is the innovation covariance. $\Xi_{k,i}^{x,\varepsilon}$ is the expectation between x_k and $\varepsilon_i (i = k, k+1)$. $P_{k|k}$ and $P_{k+1|k}$ are the state filtering error covariance and state one-step prediction error covariance, respectively. $X_{k,k}^{\mathcal{A}_s,\mathcal{A}_s} = \mathcal{E}\{\mathcal{A}_{s,k}\mathcal{A}_{s,k}^T\}$ and the matrices $\mathcal{A}_k$, $\mathcal{A}_{s,k}$ and $\mathcal{D}_k$ are defined in (48). The expectations $\Theta_{k,k|k-1}^{\omega,x}$, $\Theta_{k,k|k-1}^{v,x}$, $\Theta_{k,k|k-1}^{v,\omega}$, $\Theta_{k,k|k-1}^{\omega,\omega}$ and $\Theta_{k,k|k-1}^{v,v}$ are defined and calculated as in Lemma 3. The initial values are $\hat{x}_{0|-1} = [\bar{\vec{x}}_0^T \;\; 0]^T$, $P_{0|-1} = diag(\vec{P}_0, 0)$.

Proof: According to (5), (7), the OPT and the fact that η_k is zero mean and uncorrelated with other signals, the system state one-step predictor $\hat{x}_{k+1|k}$ can be calculated as follows:

$$\begin{aligned}\hat{x}_{k+1|k} &= \mathcal{E}\{x_{k+1}\} + \sum_{i=1}^{k}\mathcal{E}\left\{x_{k+1}\varepsilon_i^T\right\}\Lambda_i^{-1}\varepsilon_i \\ &= A_k\mathcal{E}\{x_k\} + \sum_{i=1}^{k-1}\mathcal{E}\left\{A_k x_k\varepsilon_i^T\right\}\Lambda_i^{-1}\varepsilon_i \\ &\quad + \sum_{i=1}^{k-1}\mathcal{E}\left\{B_k\omega_k\varepsilon_i^T\right\}\Lambda_i^{-1}\varepsilon_i + \mathcal{E}\left\{x_{k+1}\varepsilon_k^T\right\}\Lambda_k^{-1}\varepsilon_k \\ &= A_k\hat{x}_{k|k-1} + B_k\hat{\omega}_{k|k-1} + \Xi_{k+1,k}^{x,\varepsilon}\Lambda_k^{-1}\varepsilon_k,\end{aligned}$$

therefore, the one-step prediction error $\tilde{x}_{k+1|k}$ has the following form:

$$\begin{aligned}\tilde{x}_{k+1|k} &= (A_k - \Xi_{k+1,k}^{x,\varepsilon}\Lambda_k^{-1}\bar{C}_k)\tilde{x}_{k|k-1} + (A_{s,k}\eta_k \\ &\quad - \Xi_{k+1,k}^{x,\varepsilon}\Lambda_k^{-1}\tilde{C}_k)x_k + B_k\tilde{\omega}_{k|k-1} - \Xi_{k+1,k}^{x,\varepsilon} \\ &\quad \times \Lambda_k^{-1}\tilde{D}_k v_k - \Xi_{k+1,k}^{x,\varepsilon}\Lambda_k^{-1}\bar{D}_k\tilde{v}_{k|k-1}.\end{aligned}$$

By defining

$$\mathcal{A}_{s,k} = (A_{s,k}\eta_k - \Xi^{x,\varepsilon}_{k+1,k}\Lambda_k^{-1}\tilde{C}_k)x_k - \Xi^{x,\varepsilon}_{k+1,k}\Lambda_k^{-1}\tilde{D}_k v_k,$$
$$\mathcal{A}_k = A_k - \Xi^{x,\varepsilon}_{k+1,k}\Lambda_k^{-1}\bar{C}_k, \quad \mathcal{D}_k = \Xi^{x,\varepsilon}_{k+1,k}\Lambda_k^{-1}\bar{D}_k, \tag{48}$$

the one-step prediction error $\tilde{x}_{k+1|k}$ can be rewritten by:

$$\tilde{x}_{k+1|k} = \mathcal{A}_k\tilde{x}_{k|k-1} + \mathcal{A}_{s,k} + B_k\tilde{\omega}_{k|k-1} - \mathcal{D}_k\tilde{v}_{k|k-1}. \tag{49}$$

Furthermore, noting (49), the one-step prediction error covariance can be calculated as follows:

$$\begin{aligned} P_{k+1|k} &= \mathcal{A}_k\mathcal{E}\left\{\tilde{x}_{k|k-1}\tilde{x}^T_{k|k-1}\right\}\mathcal{A}_k^T + \mathcal{A}_k\mathcal{E}\left\{\tilde{x}_{k|k-1}\mathcal{A}^T_{s,k}\right\} \\ &+ \mathcal{A}_k\mathcal{E}\{\tilde{x}_{k|k-1}\tilde{\omega}^T_{k|k-1}\}B_k^T + \mathcal{E}\{\mathcal{A}_{s,k}\mathcal{A}^T_{s,k}\} \\ &- \mathcal{A}_k\mathcal{E}\{\tilde{x}_{k|k-1}\tilde{v}^T_{k|k-1}\}\mathcal{D}_k^T + \mathcal{E}\{\mathcal{A}_{s,k}\tilde{x}^T_{k|k-1}\}\mathcal{A}_k^T \\ &+ \mathcal{E}\{\mathcal{A}_{s,k}\tilde{\omega}_{k|k-1}\}B_k^T - \mathcal{E}\{\mathcal{A}_{s,k}\tilde{v}^T_{k|k-1}\}\mathcal{D}_k^T \\ &+ B_k\mathcal{E}\{\tilde{\omega}_{k|k-1}\tilde{x}^T_{k|k-1}\}\mathcal{A}_k^T + B_k\mathcal{E}\{\tilde{\omega}_{k|k-1}\mathcal{A}^T_{s,k}\} \\ &+ B_k\mathcal{E}\{\tilde{\omega}_{k|k-1}\tilde{\omega}^T_{k|k-1}\}B_k^T - B_k\mathcal{E}\{\tilde{\omega}_{k|k-1} \\ &\times \tilde{v}^T_{k|k-1}\}\mathcal{D}_k^T - \mathcal{D}_k\mathcal{E}\{\tilde{v}_{k|k-1}\tilde{x}^T_{k|k-1}\}\mathcal{A}_k^T - \mathcal{D}_k \\ &\times \mathcal{E}\{\tilde{v}_{k|k-1}\tilde{\omega}^T_{k|k-1}\}B_k^T - \mathcal{D}_k\mathcal{E}\{\tilde{v}_{k|k-1}\mathcal{A}^T_{s,k}\} \\ &+ \mathcal{D}_k\mathcal{E}\{\tilde{v}_{k|k-1}\tilde{v}^T_{k|k-1}\}\mathcal{D}_k^T, \end{aligned} \tag{50}$$

where the remaining expectations can be calculated by:

$$\begin{aligned} \mathcal{E}\{\tilde{\omega}_{k|k-1}\tilde{x}^T_{k|k-1}\} &= \mathcal{E}\{\omega_k\tilde{x}^T_{k|k-1}\} = \Theta^{\omega,x}_{k,k|k-1}, \\ \mathcal{E}\{\tilde{v}_{k|k-1}\tilde{x}^T_{k|k-1}\} &= \mathcal{E}\{v_k\tilde{x}^T_{k|k-1}\} = \Theta^{v,x}_{k,k|k-1}, \\ \mathcal{E}\{\tilde{v}_{k|k-1}\tilde{v}^T_{k|k-1}\} &= \mathcal{E}\{v_k\tilde{v}^T_{k|k-1}\} = \Theta^{v,v}_{k,k|k-1}, \\ \mathcal{E}\{\tilde{\omega}_{k|k-1}\tilde{\omega}^T_{k|k-1}\} &= \mathcal{E}\{\omega_k\tilde{\omega}^T_{k|k-1}\} = \Theta^{\omega,\omega}_{k,k|k-1}, \\ \mathcal{E}\{\tilde{\omega}_{k|k-1}\tilde{v}^T_{k|k-1}\} &= \mathcal{E}\{\omega_k\tilde{v}^T_{k|k-1}\} = \Theta^{\omega,v}_{k,k|k-1}, \\ \mathcal{E}\left\{\tilde{x}_{k|k-1}\mathcal{A}^T_{s,k}\right\} &= \mathcal{E}\{\tilde{v}_{k|k-1}\mathcal{A}^T_{s,k}\} = \mathcal{E}\{\tilde{\omega}_{k|k-1}\mathcal{A}^T_{s,k}\} = 0, \end{aligned} \tag{51}$$

$$\begin{aligned} X^{\mathcal{A}_s,\mathcal{A}_s}_{k,k} &= A_{s,k}X^{x,x}_{k,k}A^T_{s,k} + \sigma_k\Xi^{x,\varepsilon}_{k+1,k}\Lambda_k^{-1}\bar{D}_eR_k\bar{D}_e^T \\ &\times \Lambda_k^{-1}(\Xi^{x,\varepsilon}_{k+1,k})^T + \sigma_k\Xi^{x,\varepsilon}_{k+1,k}\Lambda_k^{-1}\bar{C}_{e,k} \\ &\times X^{x,x}_{k,k}\bar{C}^T_{e,k}\Lambda_k^{-1}(\Xi^{x,\varepsilon}_{k+1,k})^T, \end{aligned} \tag{52}$$

where Lemmas 1-4 have been applied. Substituting (51) and (52) into (50), we have (43).

It follows from (27) and the fact that matrices $\tilde{C}_k$ and $\tilde{D}_k$ are zero mean, the innovation covariance $\Lambda_k = \mathcal{E}\{\varepsilon_k\varepsilon_k^T\}$ can be calculated as follows:

$$\begin{aligned} \Lambda_k &= \mathcal{E}\left\{J_k^2\right\}\bar{C}_{e,k}\mathcal{E}\left\{x_kx_k^T\right\}\bar{C}^T_{e,k} + \bar{C}_k\mathcal{E}\left\{\tilde{x}_{k|k-1}\right. \\ &\left.\times \tilde{x}^T_{k|k-1}\right\}\bar{C}_k^T + \mathcal{E}\left\{J_k^2\right\}\bar{D}_e\mathcal{E}\{v_kv_k^T\}\bar{D}_e^T \\ &+ \bar{D}_k\mathcal{E}\{\tilde{v}_{k|k-1}\tilde{v}^T_{k|k-1}\}\bar{D}_k^T \\ &= \sigma_k\bar{C}_{e,k}X^{x,x}_{k,k}\bar{C}^T_{e,k} + \bar{C}_kP_{k|k-1}\bar{C}_k^T \\ &+ \sigma_k\bar{D}_eR_k\bar{D}_e^T + \bar{D}_k\Theta^{v,v}_{k,k|k-1}\bar{D}_k^T. \end{aligned}$$

From (27) and the fact that v_k and $\tilde{C}_k$ are zero mean uncorrelated with each, the expectation $\Xi^{x,\varepsilon}_{k,k} = \mathcal{E}\left\{x_k\varepsilon_k^T\right\}$ can be calculated as follows:

$$\begin{aligned} \Xi^{x,\varepsilon}_{k,k} &= \mathcal{E}\left\{x_k(\tilde{C}_kx_k + \bar{C}_k\tilde{x}_{k|k-1}\right. \\ &\left. + \tilde{D}_kv_k + \bar{D}_k\tilde{v}_{k|k-1})^T\right\} \\ &= P_{k|k-1}\bar{C}_k^T + (\Theta^{v,x}_{k,k|k-1})^T\bar{D}_k^T, \end{aligned}$$

where $\Theta^{v,x}_{k,k|k-1}$ can be calculated as in Lemma 3. Furthermore, the expectation $\Xi^{x,\varepsilon}_{k+1,k} = \mathcal{E}\left\{x_{k+1}\varepsilon_k^T\right\}$ can be calculated as follows:

$$\begin{aligned} \Xi^{x,\varepsilon}_{k+1,k} &= \mathcal{E}\left\{(A_kx_k + A_{s,k}\eta_kx_k + B_k\omega_k)\varepsilon_k^T\right\} \\ &= A_k\Xi^{x,\varepsilon}_{k,k} + B_k\Xi^{\omega,\varepsilon}_{k,k}, \end{aligned}$$

where $\Xi^{\omega,\varepsilon}_{k,k}$ can be calculated as in Lemma 3.

Again, by using (5) and the OPT, the recursive filter $\hat{x}_{k|k}$ can be designed as follows:

$$\begin{aligned} \hat{x}_{k|k} &= \mathcal{E}\{x_k\} + \sum_{i=1}^{k}\mathcal{E}\left\{x_k\varepsilon_i^T\right\}\Lambda_i^{-1}\varepsilon_i \\ &= \hat{x}_{k|k-1} + \Xi^{x,\varepsilon}_{k,k}\Lambda_k^{-1}\varepsilon_k, \end{aligned}$$

therefore, the filtering error $\tilde{x}_{k|k}$ is calculated as follows:

$$\tilde{x}_{k|k} = \tilde{x}_{k|k-1} - \Xi^{x,\varepsilon}_{k,k}\Lambda_k^{-1}\varepsilon_k. \tag{53}$$

Based on (53), the filtering error covariance $P_{k|k}$ can be calculated as follows:

$$\begin{aligned} P_{k|k} &= \mathcal{E}\{\tilde{x}_{k|k-1}\tilde{x}^T_{k|k-1}\} - \mathcal{E}\{\tilde{x}_{k|k-1}\varepsilon_k^T\}\Lambda_k^{-1}(\Xi^{x,\varepsilon}_{k,k})^T \\ &- \Xi^{x,\varepsilon}_{k,k}\Lambda_k^{-1}\mathcal{E}\{\varepsilon_k\tilde{x}^T_{k|k-1}\} + \Xi^{x,\varepsilon}_{k,k}\Lambda_k^{-1}(\Xi^{x,\varepsilon}_{k,k})^T, \end{aligned}$$

where the expectation $\mathcal{E}\{\tilde{x}_{k|k-1}\tilde{x}^T_{k|k-1}\} = P_{k|k-1}$ and the expectation $\mathcal{E}\{\tilde{x}_{k|k-1}\varepsilon_k^T\}$ can be calculated as follows:

$$\mathcal{E}\{\tilde{x}_{k|k-1}\varepsilon_k^T\} = \mathcal{E}\{x_k\varepsilon_k^T\} - \mathcal{E}\left\{\left\{\sum_{i=1}^{k-1}\mathcal{E}\{x_k\varepsilon_i^T\}\Lambda_i^{-1}\varepsilon_i\right\}\varepsilon_k^T\right\}.$$

The fact that ε_k is uncorrelated with $\varepsilon_j, j \neq k$, we have

$$\mathcal{E}\{\tilde{x}_{k|k-1}\varepsilon_k^T\} = \mathcal{E}\{x_k\varepsilon_k^T\} - 0 = \Xi^{x,\varepsilon}_{k,k},$$

therefore, the error covariance $P_{k|k}$ can be rewritten by:

$$P_{k|k} = P_{k|k-1} - \Xi^{x,\varepsilon}_{k,k}\Lambda_k^{-1}(\Xi^{x,\varepsilon}_{k,k})^T,$$

which completes the proof of theorem 1.

We are now in a position to proceed with the design of predictor and smoother for the given system.

B. ROBUST RECURSIVE PREDICTOR

Theorem 2: For the given system (5)-(6), we have the following predictor:

$$\begin{cases} \mathbf{1 < N \leq t:} \\ \hat{x}_{k+N|k} = A_{k+N-1}\hat{x}_{k+N-1|k} + B_{k+N-1}\hat{\omega}_{k+N-1|k}, \\ \hat{\omega}_{k+N-1|k} = \sum_{i=0}^{t-N} \Xi^{\omega,\varepsilon}_{k+N-1,k-i}\Lambda^{-1}_{k-i}\varepsilon_{k-i}, \\ P_{k+N|k} = A_{k+N-1}P_{k+N-1|k}A^T_{k+N-1} \\ \qquad + A_{s,k+N-1}X^{x,x}_{k+N-1,k+N-1} \\ \qquad \times A^T_{s,k+N-1} + A_{k+N-1} \\ \qquad \times \Theta^{x,\omega}_{k+N-1,k+N-1|k}B^T_{k+N-1} \\ \qquad + B_{k+N-1}(\Theta^{x,\omega}_{k+N-1,k+N-1|k})^T \\ \qquad \times A^T_{k+N-1} + B_{k+N-1} \\ \qquad \times \Theta^{\omega,\omega}_{k+N-1,k+N-1}B^T_{k+N-1}, \\ \Theta^{\omega,\omega}_{k+N-1,k+N-1|k} = Q^{\omega}_{k+N-1,k+N-1} \\ \qquad - \sum_{i=0}^{t-N} \Xi^{\omega,\varepsilon}_{k+N-1,k-i} \\ \qquad \times \Lambda^{-1}_{k-i}(\Xi^{\omega,\varepsilon}_{k+N-1,k-i})^T, \\ \Theta^{x,\omega}_{k+N-1,k+N-1|k} = \prod_{n=2}^{t} A_{k+N-n}B_{k+N-t-1} \\ \qquad \times Q_{k+N-t-1,k+N-1} \\ \qquad - \sum_{i=0}^{t-N} \Xi^{x,\varepsilon}_{k+N-1,k-i} \\ \qquad \times \Lambda^{-1}_{k-i}(\Xi^{\omega,\varepsilon}_{k+N-1,k-i})^T, \\ \Xi^{x,\varepsilon}_{k+N-1,k-i} = \Theta^{x,x}_{k+N-1,k-i|k-i-1}\bar{C}^T_{k-i} \\ \qquad + \Theta^{x,v}_{k+N-1,k-i|k-i-1}\bar{D}^T_{k-i}, \\ \Theta^{x,x}_{k+N-1,k-i|k-i-1} = \prod_{n=2}^{N+i} A_{k+N-n} \\ \qquad \times P_{k-i|k-i-1} \\ \qquad + \sum_{n=2}^{N+i} \urcorner_{n-1}B_{k+N-n} \\ \qquad \times \Theta^{\omega,x}_{k+N-n,k-i|k-i-1}, \\ \Theta^{x,v}_{k+N-1,k-i|k-i-1} = \prod_{n=2}^{N+i} A_{k+N-n} \\ \qquad \times \Theta^{x,v}_{k-i,k-i|k-i-1} \\ \qquad + \sum_{n=2}^{N+i} \urcorner_{n-1}B_{k+N-n} \\ \qquad \times \Theta^{\omega,v}_{k+N-n,k-i|k-i-1}, \end{cases} \tag{54}$$

$$\begin{cases} \mathbf{N > t:} \\ \hat{x}_{k+N|k} = A_{k+N-1}\hat{x}_{k+N-1|k}, \\ P_{k+N|k} = A_{k+N-1}P_{k+N-1}A^T_{k+N-1} \\ \qquad + A_{s,k+N-1}X^{x,x}_{k+N-1,k+N-1} \\ \qquad \times A^T_{s,k+N-1} + B_{k+N-1}\, Q_{k+N-1} \\ \qquad \times B^T_{k+N-1} + A_{k+N-1} \\ \qquad \times \Theta^{x,\omega}_{k+N-1|k,k+N-1}B^T_{k+N-1} \\ \qquad + B_{k+N-1}(\Theta^{x,\omega}_{k+N-1|k,k+N-1})^T \\ \qquad \times A^T_{k+N-1}, \\ \Theta^{x,\omega}_{k+N-1|k,k+N-1} = X^{x,\omega}_{k+N-1,k+N-1}, \end{cases} \tag{55}$$

where the initial values are given in Theorem 1 and Lemmas 1-4. The notation $\urcorner_n$ is defined in (58).

Proof: From (5) and the OPT, the N-step state predictor $\hat{x}_{k+N|k}$ can be calculated as follows:

$$\begin{aligned} \hat{x}_{k+N|k} &= \mathcal{E}\{x_{k+N}\} + \sum_{i=1}^{k} \mathcal{E}\{x_{k+N}\varepsilon_i^T\}\Lambda_i^{-1}\varepsilon_i \\ &= A_{k+N-1}\hat{x}_{k+N-1|k} + B_{k+N-1}\hat{\omega}_{k+N-1|k}, \end{aligned} \tag{56}$$

where the process noise predictor $\hat{\omega}_{k+N-1|k}$ can be calculated as follows:

$$\hat{\omega}_{k+N-1|k} = \sum_{i=0}^{k-1} \mathcal{E}\{\omega_{k+N-1}\varepsilon^T_{k-i}\}\Lambda^{-1}_{k-i}\varepsilon_{k-i}.$$

It can be easily seen that if $N > t$, then the expectation $\mathcal{E}\{\omega_{k+N-1}\varepsilon^T_{k-i}\} = 0$, furthermore, we have $\hat{\omega}_{k+N-1|k} = 0$. Therefore, the design of the proposed N-step predictor can be divided into two different parts: One is $N > t$ and the other is $1 < N \leq t$.

1) $1 < N \leq t$: Taking into account the fact that the process noise is t-step discrete autocorrelated across time, we have from (5) that

$$\begin{aligned} \hat{\omega}_{k+N-1|k} &= \sum_{i=0}^{k-1} \mathcal{E}\{\omega_{k+N-1}\varepsilon^T_{k-i}\}\Lambda^{-1}_{k-i}\varepsilon_{k-i} \\ &= \sum_{i=0}^{t-N} \Xi^{\omega,\varepsilon}_{k+N-1,k-i}\Lambda^{-1}_{k-i}\varepsilon_{k-i}, \end{aligned}$$

therefore, the $N-1$ step process noise prediction error $\tilde{\omega}_{k+N-1|k}$ can be expressed as follows:

$$\tilde{\omega}_{k+N-1|k} = \omega_{k+N-1} - \sum_{i=0}^{t-N} \Xi^{\omega,\varepsilon}_{k+N-1,k-i}\Lambda^{-1}_{k-i}\varepsilon_{k-i}.$$

It follows from (5) and (56) that the N-step state prediction error $\tilde{x}_{k+N|k}$ can be calculated as follows:

$$\begin{aligned} \tilde{x}_{k+N|k} = A_{k+N-1}\tilde{x}_{k+N-1|k} &+ A_{s,k+N-1}\eta_{k+N-1} \\ &\times x_{k+N-1} + B_{k+N-1}\tilde{\omega}_{k+N-1|k}, \end{aligned}$$

and then, the N-step state prediction error covariance $P_{k+N|k}$ can be calculated as follows:

$$\begin{aligned} P_{k+N|k} = A_{k+N-1}P_{k+N-1|k}A^T_{k+N-1} &+ A_{s,k+N-1} \\ \times X^{x,x}_{k+N-1,k+N-1}A^T_{s,k+N-1} &+ A_{k+N-1} \\ \times \mathcal{E}\{\tilde{x}_{k+N-1|k}\tilde{\omega}^T_{k+N-1|k}\}&B^T_{k+N-1} \\ + B_{k+N-1}\mathcal{E}\{\tilde{\omega}_{k+N-1|k}&\tilde{x}^T_{k+N-1|k}\} \\ \times A^T_{k+N-1} + B_{k+N-1}&\mathcal{E}\{\tilde{\omega}_{k+N-1|k} \\ \times \tilde{\omega}^T_{k+N-1|k}\}&B^T_{k+N-1}. \end{aligned} \tag{57}$$

Our next step is to calculate the remaining expectations in (57). From Lemma 4, we have

$$\begin{aligned} \mathcal{E}\{\tilde{\omega}_{k+N-1|k}\tilde{\omega}^T_{k+N-1|k}\} &= \mathcal{E}\{\omega_{k+N-1}\tilde{\omega}^T_{k+N-1|k}\} \\ &= \Theta^{\omega,\omega}_{k+N-1,k+N-1|k}, \end{aligned}$$

and the expectation $\Theta^{\omega,\omega}_{k+N-1,k+N-1|k}$ can be calculated as follows:

$$\begin{aligned}\Theta^{\omega,\omega}_{k+N-1,k+N-1|k} &= \mathcal{E}\{\omega_{k+N-1}\omega^T_{k+N-1}\} \\ &\quad - \mathcal{E}\{\omega_{k+N-1}\hat{\omega}^T_{k+N-1|k}\} \\ &= Q^{\omega}_{k+N-1,k+N-1} - \sum_{i=0}^{t-N} \Xi^{\omega,\varepsilon}_{k+N-1,k-i} \\ &\quad \times \Lambda^{-1}_{k-i}(\Xi^{\omega,\varepsilon}_{k+N-1,k-i})^T,\end{aligned}$$

where the fact that the process noise is t-step discrete autocorrelated across time has been applied.

By introducing the notation $\daleth_n$

$$\daleth_n = \begin{cases} \prod_{m=2}^{n} A_{k+N-m}, & n > 1, \\ I, & n = 1, \end{cases} \tag{58}$$

the state x_{k+N-1} can be rewritten as follows:

$$\begin{aligned}x_{k+N-1} &= A_{k+N-2}x_{k+N-2} + A_{s,k+N-2} \\ &\quad \times \eta_{k+N-2}x_{k+N-2} + B_{k+N-2}\omega_{k+N-2} \\ &= \prod_{n=2}^{t} A_{k+N-n}x_{k+N-t} + \sum_{n=2}^{t} \daleth_{n-1} \\ &\quad \times (A_{s,k+N-n}\eta_{k+N-n}x_{k+N-n} \\ &\quad + B_{k+N-n}\omega_{k+N-n}),\end{aligned}$$

therefore, the expectation $\mathcal{E}\{\tilde{x}_{k+N-1|k}\tilde{\omega}^T_{k+N-1|k}\} = \mathcal{E}\{x_{k+N-1}\tilde{\omega}^T_{k+N-1|k}\} = \Theta^{x,\omega}_{k+N-1,k+N-1|k}$ can be calculated as follows:

$$\begin{aligned}\Theta^{x,\omega}_{k+N-1,k+N-1|k} &= \mathcal{E}\{x_{k+N-1}\omega^T_{k+N-1}\} \\ &\quad - \mathcal{E}\{x_{k+N-1}\hat{\omega}^T_{k+N-1|k}\} \\ &= \prod_{n=2}^{t} A_{k+N-n}B_{k+N-t-1} \\ &\quad \times Q_{k+N-t-1,k+N-1} \\ &\quad - \sum_{i=0}^{t-N} \Xi^{x,\varepsilon}_{k+N-1,k-i} \\ &\quad \times \Lambda^{-1}_{k-i}(\Xi^{\omega,\varepsilon}_{k+N-1,k-i})^T,\end{aligned}$$

where the second equality holds since the process noise is t-step discrete autocorrelated across time and η_k is uncorrelated with other signals. Taking (27) into consideration, the remaining expectation $\Xi^{x,\varepsilon}_{k+N-1,k-i} = \mathcal{E}\{x_{k+N-1}\varepsilon^T_{k-i}\}$ can be calculated as follows:

$$\begin{aligned}\Xi^{x,\varepsilon}_{k+N-1,k-i} &= \mathcal{E}\left\{x_{k+N-1}\left(\tilde{C}_{k-i}x_{k-i} + \bar{C}_{k-i}\tilde{x}_{k-i|k-i-1}\right.\right. \\ &\quad \left.\left. + \tilde{D}_{k-i}v_{k-i} + \bar{D}_{k-i}\tilde{v}_{k-i|k-i-1}\right)^T\right\} \\ &= \mathcal{E}\{x_{k+N-1}x^T_{k-i}\}\mathcal{E}\{\tilde{C}^T_{k-i}\} \\ &\quad + \mathcal{E}\{x_{k+N-1}\tilde{x}^T_{k-i|k-i-1}\}\bar{C}^T_{k-i} \\ &\quad + \mathcal{E}\{x_{k+N-1}v^T_{k-i}\}\mathcal{E}\{\tilde{D}^T_{k-i}\} \\ &\quad + \mathcal{E}\{x_{k+N-1}\tilde{v}^T_{k-i|k-i-1}\}\bar{D}^T_{k-i},\end{aligned}$$

where the matrices $\tilde{C}_{k-i}$ and $\tilde{D}_{k-i}$ are zero mean, therefore, the expectation $\Xi^{x,\varepsilon}_{k+N-1,k-i}$ can be rewritten as follows:

$$\begin{aligned}\Xi^{x,\varepsilon}_{k+N-1,k-i} &= \Theta^{x,x}_{k+N-1,k-i|k-i-1}\bar{C}^T_{k-i} \\ &\quad + \Theta^{x,v}_{k+N-1,k-i|k-i-1}\bar{D}^T_{k-i},\end{aligned}$$

where $\Theta^{x,x}_{k+N-1,k-i|k-i-1} = \mathcal{E}\{x_{k+N-1}\tilde{x}^T_{k-i|k-i-1}\}$ and $\Theta^{x,v}_{k+N-1,k-i|k-i-1} = \mathcal{E}\{x_{k+N-1}\tilde{v}^T_{k-i|k-i-1}\}$ can be calculated as follows:

$$\begin{aligned}\Theta^{x,x}_{k+N-1,k-i|k-i-1} &= \prod_{n=2}^{N+i} A_{k+N-n}\mathcal{E}\left\{x_{k-i}\tilde{x}^T_{k-i|k-i-1}\right\} \\ &\quad + \sum_{n=2}^{N+i} \daleth_{n-1}B_{k+N-n} \\ &\quad \times \mathcal{E}\{\omega_{k+N-n}\tilde{x}^T_{k-i|k-i-1}\} \\ &= \prod_{n=2}^{N+i} A_{k+N-n}P_{k-i|k-i-1} \\ &\quad + \sum_{n=2}^{N+i} \daleth_{n-1}B_{k+N-n} \\ &\quad \times \Theta^{\omega,x}_{k+N-n,k-i|k-i-1}, \\ \Theta^{x,v}_{k+N-1,k-i|k-i-1} &= \prod_{n=2}^{N+i} A_{k+N-n}\mathcal{E}\left\{x_{k-i}\tilde{v}^T_{k-i|k-i-1}\right\} \\ &\quad + \sum_{n=2}^{N+i} \daleth_{n-1}B_{k+N-n} \\ &\quad \times \mathcal{E}\{\omega_{k+N-n}\tilde{v}^T_{k-i|k-i-1}\} \\ &= \prod_{n=2}^{N+i} A_{k+N-n}\Theta^{x,v}_{k-i,k-i|k-i-1} \\ &\quad + \sum_{n=2}^{N+i} \daleth_{n-1}B_{k+N-n} \\ &\quad \times \Theta^{\omega,v}_{k+N-n,k-i|k-i-1}.\end{aligned}$$

2) $N > t$: From (7) and the OPT, we have $\hat{\omega}_{k+N-1} = 0$, therefore, the N-step state predictor can be calculated as follows:

$$\hat{x}_{k+N|k} = A_{k+N-1}\hat{x}_{k+N-1|k},$$

and then, the N-step state prediction error $\tilde{x}_{k+N|k} = \hat{x}_{k+N|k} - x_{k+N}$ has the following expression:

$$\begin{aligned}\tilde{x}_{k+N|k} &= A_{k+N-1}\tilde{x}_{k+N-1|k} + A_{s,k+N-1}\eta_{k+N-1} \\ &\quad \times x_{k+N-1} + B_{k+N-1}\omega_{k+N-1}.\end{aligned} \tag{59}$$

It implies from (59) and Lemma 2 that the N-step state prediction error covariance $P_{k+N|k}$ can be calculated as follows:

$$\begin{aligned}P_{k+N|k} &= A_{k+N-1}P_{k+N-1}A^T_{k+N-1} + A_{s,k+N-1} \\ &\quad \times X^{x,x}_{k+N-1,k+N-1}A^T_{s,k+N-1} + B_{k+N-1} \\ &\quad \times Q_{k+N-1}B^T_{k+N-1} + A_{k+N-1}\end{aligned}$$

$$\times \Theta^{x,\omega}_{k+N-1|k,k+N-1}B^T_{k+N-1} + B_{k+N-1} \times (\Theta^{x,\omega}_{k+N-1|k,k+N-1})^T A^T_{k+N-1},$$

where $\Theta^{x,\omega}_{k+N-1|k,k+N-1} = \mathcal{E}\{\tilde{x}_{k+N-1|k}\omega^T_{k+N-1}\}$ can be calculated as follows:

$$\begin{aligned}\Theta^{x,\omega}_{k+N-1|k,k+N-1} &= \mathcal{E}\{x_{k+N-1}\omega^T_{k+N-1}\} - \mathcal{E}\{\hat{x}_{k+N-1|k}\omega^T_{k+N-1}\}\\ &= \prod_{n=1}^{t-1} A_{k+N-1-(n-1)}\mathcal{E}\{x_{k+N-t} \times \omega_{k+N-1}\} - \mathcal{E}\{\hat{x}_{k+N-1|k}\omega^T_{k+N-1}\}\\ &= X^{x,\omega}_{k+N-1,k+N-1} - \sum_{i=1}^{k}\Xi^{x,\varepsilon}_{k+N-1,i} \times \Lambda_i^{-1}\mathcal{E}\{\varepsilon_i\omega^T_{k+N-1}\},\end{aligned}$$

where Lemma 1 has been applied. Since $k+N-1-k = N-1 \geq t$, therefore, $\mathcal{E}\{\varepsilon_i\omega^T_{k+N-1}\} = 0$, and then, we have $\Theta^{x,\omega}_{k+N-1|k,k+N-1} = X^{x,\omega}_{k+N-1,k+N-1}$ which completes the proof of the Theorem 2.

C. ROBUST RECURSIVE SMOOTHER

Theorem 3: For the addressed system (5)-(6), the N-step ($N > 0$) fixed-lag robust recursive smoother can be calculated as follows:

$$\hat{x}_{k|k+N} = \hat{x}_{k|k+N-1} + \Xi^{x,\varepsilon}_{k,k+N}\Lambda^{-1}_{k+N}\varepsilon_{k+N}, \quad (60)$$

$$P_{k|k+N} = P_{k|k+N-1} - \Xi^{x,\varepsilon}_{k,k+N}\Lambda^{-1}_{k+N}(\Xi^{x,\varepsilon}_{k,k+N})^T, \quad (61)$$

$$\Xi^{x,\varepsilon}_{k,k+N} = \Theta^{x,x}_{k,k+N|k+N-1}\bar{C}^T_{k+N} + \Theta^{x,v}_{k,k+N|k+N-1}\bar{D}^T_{k+N}, \quad (62)$$

$$\Theta^{x,x}_{k,k|k-1} = P_{k|k-1}, \quad (63)$$

$$\Theta^{x,x}_{k,k+N|k+N-1} = \Theta^{x,x}_{k,k+N-1|k+N-2}\mathcal{A}^T_{k+N-1} + \Theta^{x,\omega}_{k,k+N-1|k+N-2}B^T_{k+N-1} - \Theta^{x,v}_{k,k+N-1|k+N-2}\mathcal{D}^T_{k+N-1}, \quad (64)$$

$$\Theta^{x,v}_{k,k+N|k+N-1} = -\sum_{i=1}^{k+N-1}\Xi^{x,\varepsilon}_{k,i}\Lambda_i^{-1}(\Xi^{v,\varepsilon}_{k+N,i})^T, \quad (65)$$

$$\Theta^{x,\omega}_{k,k+N-1|k+N-2} = X^{x,\omega}_{k,k+N-1} - \sum_{i=1}^{k+N-2}\Xi^{x,\varepsilon}_{k,i} \times \Lambda_i^{-1}(\Xi^{\omega,\varepsilon}_{k+N-1,i})^T, \quad (66)$$

where the matrices $\mathcal{A}_{k+N-1}$ and $\mathcal{D}_{k+N-1}$ are defined in (48). The initial values are given by Theorem 1 and Lemmas 1-4.

Proof: Applying the OPT, the N-step fixed-lag robust recursive smoother is calculated as follows:

$$\begin{aligned}\hat{x}_{k|k+N} &= \mathcal{E}\{x_k\} + \sum_{i=1}^{k+N}\mathcal{E}\{x_k\varepsilon_i^T\}\Lambda_i^{-1}\varepsilon_i\\ &= \hat{x}_{k|k+N-1} + \Xi^{x,\varepsilon}_{k,k+N}\Lambda^{-1}_{k+N}\varepsilon_{k+N}, \quad (67)\end{aligned}$$

where the expectation $\Xi^{x,\varepsilon}_{k,k+N} = \mathcal{E}\{x_k\varepsilon^T_{k+N}\}$ can be calculated as follows:

$$\begin{aligned}\Xi^{x,\varepsilon}_{k,k+N} &= \mathcal{E}\Big\{x_k(\tilde{C}_{k+N}x_{k+N} + \bar{C}_{k+N}\tilde{x}_{k+N|k+N-1} + \tilde{D}_{k+N}v_{k+N} + \bar{D}_{k+N}\tilde{v}_{k+N|k+N-1})^T\Big\}\\ &= \Theta^{x,x}_{k,k+N|k+N-1}\bar{C}^T_{k+N} + \Theta^{x,v}_{k,k+N|k+N-1}\bar{D}^T_{k+N}, \quad (68)\end{aligned}$$

where the last equality holds since the matrices $\tilde{C}_{k+N}$ and $\tilde{D}_{k+N}$ are zero mean and uncorrelated with the system state. It follows from the OPT that the expectations $\Theta^{x,v}_{k,k+N|k+N-1} = \mathcal{E}\{x_k\tilde{v}^T_{k+N|k+N-1}\}$ and $\Theta^{x,x}_{k,k+N|k+N-1} = \mathcal{E}\{x_k\tilde{x}^T_{k+N|k+N-1}\}$ in (68) can be calculated as follows:

$$\begin{aligned}\Theta^{x,v}_{k,k+N|k+N-1} &= \mathcal{E}\{x_k(v_{k+N} - \hat{v}_{k+N|k+N-1})^T\}\\ &= -\sum_{i=1}^{k+N-1}\Xi^{x,\varepsilon}_{k,i}\Lambda_i^{-1}(\Xi^{v,\varepsilon}_{k+N,i})^T,\\ \Theta^{x,x}_{k,k+N|k+N-1} &= \mathcal{E}\{x_k(\mathcal{A}_{k+N-1}\tilde{x}_{k+N-1|k+N-2} + \mathcal{A}_{s,k+N-1} + B_{k+N-1}\tilde{\omega}_{k+N-1|k+N-2} - \mathcal{D}_{k+N-1}\tilde{v}_{k+N-1|k+N-2})^T\}\\ &= \Theta^{x,x}_{k,k+N-1|k+N-2}\mathcal{A}^T_{k+N-1} + \Theta^{x,\omega}_{k,k+N-1|k+N-2}B^T_{k+N-1} - \Theta^{x,v}_{k,k+N-1|k+N-2}\mathcal{D}^T_{k+N-1}, \quad (69)\end{aligned}$$

where the expectation $\Theta^{x,\omega}_{k,k+N-1|k+N-2}$ in (69) can be calculated as follows:

$$\begin{aligned}\Theta^{x,\omega}_{k,k+N-1|k+N-2} &= \mathcal{E}\{x_k\omega^T_{k+N-1}\} - \mathcal{E}\{x_k\hat{\omega}^T_{k+N-1|k+N-2}\}\\ &= X^{x,\omega}_{k,k+N-1} - \sum_{i=1}^{k+N-2}\Xi^{x,\varepsilon}_{k,i}\Lambda_i^{-1}(\Xi^{\omega,\varepsilon}_{k+N-1,i})^T, \quad (70)\end{aligned}$$

where Lemma 1 has been applied in (70). In addition, applying Lemma 4, we have

$$\Theta^{x,x}_{k,k|k-1} = \mathcal{E}\{x_k\tilde{x}^T_{k|k-1}\} = \mathcal{E}\{\tilde{x}_{k|k-1}\tilde{x}^T_{k|k-1}\} = P_{k|k-1}.$$

From (67), the smoother error can be obtained as follows:

$$\begin{aligned}\tilde{x}_{k|k+N} &= x_k - \hat{x}_{k|k+N-1} - \Xi^{x,\varepsilon}_{k,k+N}\Lambda^{-1}_{k+N}\varepsilon_{k+N}\\ &= \tilde{x}_{k|k+N-1} - \Xi^{x,\varepsilon}_{k,k+N}\Lambda^{-1}_{k+N}\varepsilon_{k+N},\end{aligned}$$

therefore, the smoother error covariance can be obtained by:

$$\begin{aligned}P_{k|k+N} &= P_{k|k+N-1} - \mathcal{E}\{\tilde{x}_{k|k+N-1}\varepsilon^T_{k+N}\}\Lambda_{k+N} \times (\Xi^{x,\varepsilon}_{k,k+N})^T - \Xi^{x,\varepsilon}_{k,k+N}\Lambda^{-1}_{k+N}\mathcal{E}\{\varepsilon_{k+N} \times \tilde{x}^T_{k|k+N-1}\} + \Xi^{x,\varepsilon}_{k,k+N}\Lambda^{-1}_{k+N}(\Xi^{x,\varepsilon}_{k,k+N})^T\\ &= P_{k|k+N-1} - \Xi^{x,\varepsilon}_{k,k+N}\Lambda^{-1}_{k+N}(\Xi^{x,\varepsilon}_{k,k+N})^T,\end{aligned}$$

which completes the proof of the Theorem 3.

Remark 6: A seemingly natural way of handling the delayed process noises and measurement noises is to augment the system states. However, when t and h are very large, such a state augmentation approach gives rise to significant increase of the system dimension, which would inevitably lead to heavy computational burden. Without resorting to state augmentation, in our current work, by combining the noises at present time and the delayed noises into a whole one, the delayed noises are transformed to be the discrete autocorrelated noises across time. At last, by using an innovation analysis approach and the OPT, the desired recursive filter, recursive predictor and recursive smoother are obtained in Theorems 1-3, respectively.

In this manuscript, the measurement delay is assumed to be at most one-step delay. However, in practical applications, there is also the phenomenon of multi-step delay. In addition, for multi-sensor fusion system, the measurement delay of each sensor is also different. These problems mentioned above are exactly the new research topics that the author will carry out on the basis of this paper in the future.

IV. AN ILLUSTRATIVE EXAMPLE

An example is provided to illustrate the effectiveness of our approaches in this manuscript. Let us consider the following uncertain system with delayed measurements and noises:

$$\begin{aligned}\vec{x}_{k+1} &= \left(\begin{bmatrix}0.5 & -0.8\\ 0.4 & 0.5\end{bmatrix} + \begin{bmatrix}0.1 & 0\\ 0 & 0.1\end{bmatrix}\eta_k\right)\vec{x}_k\\ &\quad + \begin{bmatrix}0.6\\ -1\end{bmatrix}\omega_k + \begin{bmatrix}0.2\\ 0.3\end{bmatrix}\omega_{k-t},\\ \vec{y}_k &= \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}x_k + \vec{v}_k + \vec{v}_{k-h},\\ y_k &= (1-\lambda_k)\vec{y}_k + \lambda_k\vec{y}_{k-1}, \end{aligned} \tag{71}$$

where $\vec{x}_k = [\vec{x}_{1,k}^T \quad \vec{x}_{2,k}^T]^T \in \mathbb{R}^2$ is the system state to be estimated. The scalars t and h are the time delay of the process noise and the measurement noise, respectively. The vectors $\eta_k \in \mathbb{R}$, $\omega_k \in \mathbb{R}$ and $\vec{v}_k \in \mathbb{R}^2$ are zero-mean Gaussian white noises with covariances 1, 1 and I, respectively. The variables $\lambda_k \in \mathbb{R}$ is a binary switching sequence taking values on 1 with $Prob\{\lambda_k = 1\} = \mathcal{E}\{\lambda_k\} = \beta_k = 0.15$. Our objective is to find robust recursive filter $\hat{\vec{x}}_{i,k|k}$, predictor $\hat{\vec{x}}_{i,k+2|k}$ and smoother $\hat{\vec{x}}_{i,k|k+1}$, $i = 1, 2$, and to give a comparison of their accuracies.

In the simulation, the initial value $\vec{x}_0$ has mean $\bar{\vec{x}}_0 = [0 \ \ 0]$ and the covariance $\vec{P}_0 = diag(2, 1)$. The steps of the delayed process noise and measurement noise are set as $t = 2$ and $h = 3$, respectively. Let MSE1 denote the mean square error for the estimation of $\vec{x}_{1,k}$, i.e., $(1/K)\sum_{k=1}^{K}\{\vec{x}_{1,k} - \hat{\vec{x}}_{1,k|k}\}^2$, where K is the number of the samples. Similarly, MSE2 is the mean square error for the estimation of $\vec{x}_{2,k}$, i.e., $(1/K)\sum_{k=1}^{K}\{\vec{x}_{2,k} - \hat{\vec{x}}_{2,k|k}\}^2$. Figures 1-8 are simulation results.

From Figures 1-8, we can see that, 1) the proposed robust recursive filter, predictor and smoother have good performances for the uncertain system (71), this is due to the fact that efforts have been made to compensate the

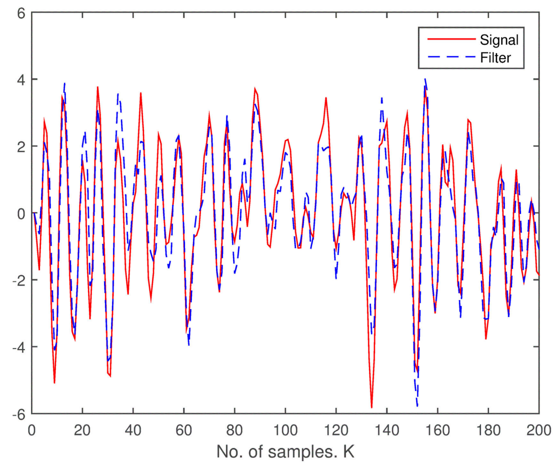


FIGURE 1. The signal of $\vec{x}_{1,k}$ and filter $\hat{\vec{x}}_{1,k|k}$.

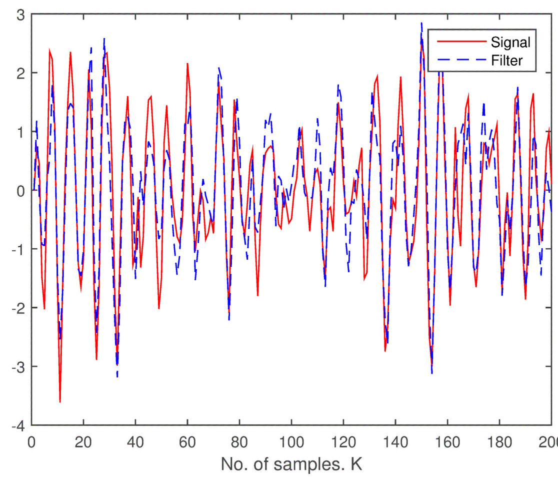


FIGURE 2. The signal of $\vec{x}_{2,k}$ and filter $\hat{\vec{x}}_{2,k|k}$.

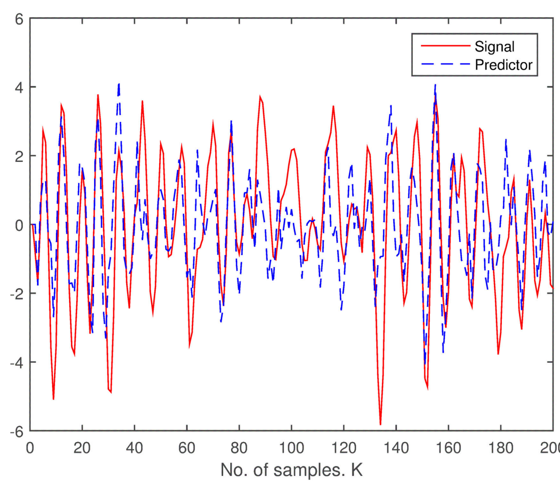


FIGURE 3. The signal of $\vec{x}_{1,k}$ and predictor $\hat{\vec{x}}_{1,k+2|k}$.

stochastic uncertainty, the randomly delayed measurements and the deterministic delayed process noise and measurement noise; 2) the predictor has the worst performance and the

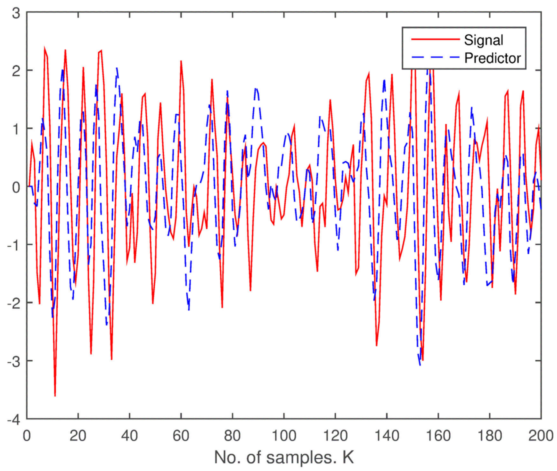


FIGURE 4. The signal of $\vec{x}_{2,k}$ and predictor $\hat{\vec{x}}_{2,k+2|k}$.

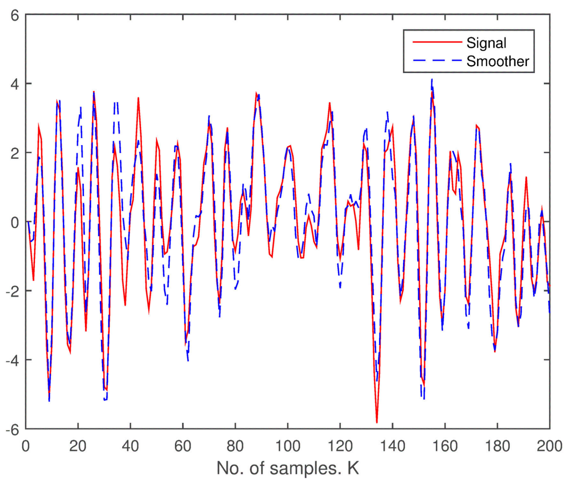


FIGURE 5. The signal of $\vec{x}_{1,k}$ and smoother $\hat{\vec{x}}_{1,k|k+1}$.

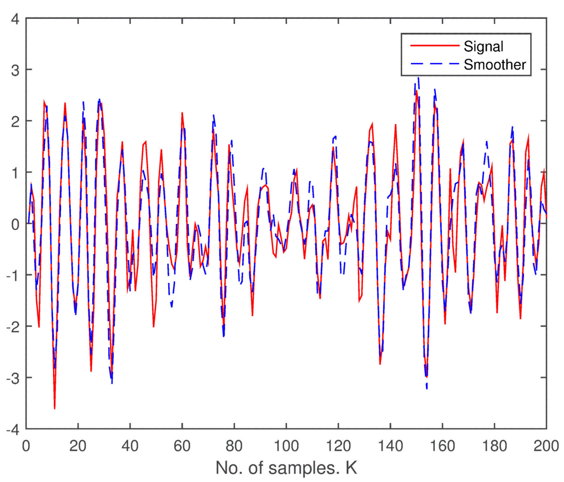


FIGURE 6. The signal of $\vec{x}_{2,k}$ and smoother $\hat{\vec{x}}_{2,k|k+1}$.

smoother has the best performance, this is natural because that the most information is used in the smoother and the least information is used in the predictor.

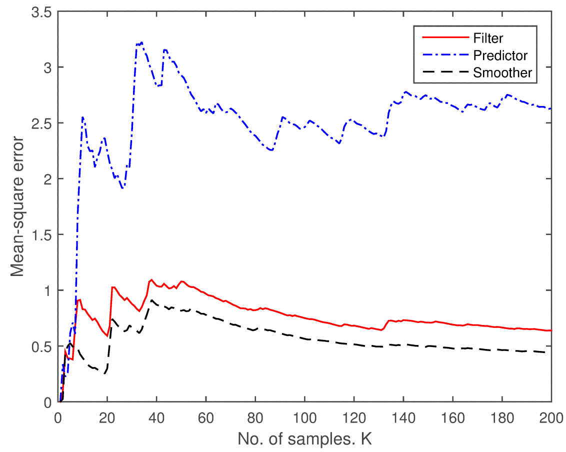


FIGURE 7. MSE1.

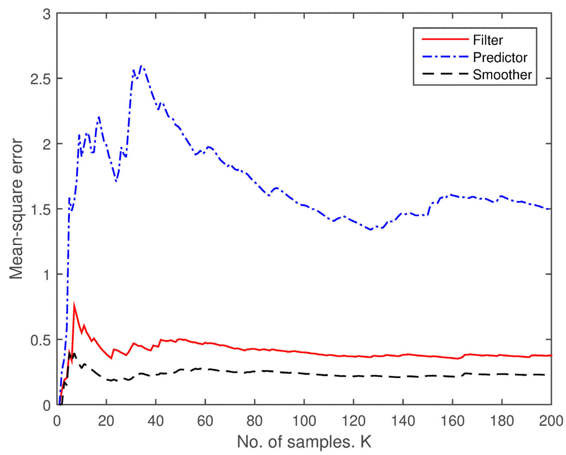


FIGURE 8. MSE2.

V. CONCLUSION

In this paper, we have investigated the robust recursive estimation problem for a class of uncertain systems with randomly delayed measurements and deterministic delayed process noises and measurement noises. The dynamic system under consideration is subject to stochastic uncertainty. The delay phenomenon of the measurements is randomly and the delay rate is described as a binary switching sequence obeying a conditional probability distribution. The process noise is assumed to be t-step time delay and the measurement noise is assumed to be h-step time delay. By combining the noise at present time and the delayed noise into a whole one, the original system is transformed to be a stochastic parameter uncertain system with discrete autocorrelated noise across time. As shown in Remark 2 and Remark 4, the discrete autocorrelated noise across time is quite different from the continuous autocorrelated noises, however, perhaps fortunately, the orthogonal projection theorem and an innovation analysis approach can be used to treat this complex case, and the obtained robust recursive estimators including filter, predictor and smoother are optimal in the linear minimum variance sense. A simulation example has been exploited to show the effectiveness of the proposed approaches.

REFERENCES

[1] R. Yang, P. Shi, and G.-P. Liu, "Filtering for discrete-time networked nonlinear systems with mixed random delays and packet dropouts," *IEEE Trans. Autom. Control.*, vol. 56, no. 11, pp. 2655–2660, Nov. 2011.

[2] E. Song, Y. Zhu, J. Zhou, and Z. You, "Optimal Kalman filtering fusion with cross-correlated sensor noises," *Automatica*, vol. 43, no. 8, pp. 1450–1456, Aug. 2007.

[3] E. Song, Y. Zhu, and Z. You, "The Kalman type recursive state estimator with a finite-step correlated process noises," in *Proc. IEEE Int. Conf. Automat. Logistics*, Sep. 2008, pp. 196–200.

[4] H. Gao, X. Meng, and T. Chen, "A parameter-dependent approach to robust H_∞ filtering for time-delay systems," *IEEE Trans. Autom. Control*, vol. 53, no. 10, pp. 2420–2425, 2008.

[5] H. Gao, Y. Zhao, J. Lam, and K. Chen, "H_∞ fuzzy filtering of nonlinear systems with intermittent measurements," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 2, pp. 291–300, 2009.

[6] H. Liu, X. Meng, Z. Chen, S. Stephenson, and P. Peltola, "A closed-loop EKF and multi-failure diagnosis approach for cooperative GNSS positioning," *GPS Solutions*, vol. 20, no. 4, pp. 795–805, Oct. 2016.

[7] A. Sayed, "A framework for state-space estimation with uncertain models," *IEEE Trans. Autom. Control.*, vol. 46, no. 7, pp. 998–1013, Jul. 2001.

[8] A. Subramanian and A. Sayed, "A robust minimum-variance filter for time varying uncertain discrete-time systems," in *Proc. Amer. Control Conf.*, Jan. 2004, pp. 1885–1889.

[9] J. Feng, Z. Wang, and M. Zeng, "Recursive robust filtering with finite-step correlated process noises and missing measurements," *Circuits Syst. Signal Process.*, vol. 30, no. 6, pp. 1355–1368, Dec. 2011.

[10] L. Zhang, P. Shi, C. Wang, and E.-K. Boukas, "Robust l_2–l_∞ filtering for switched linear discrete time-delay systems with polytopic uncertainties," *IET Control Theory Appl.*, vol. 1, no. 3, pp. 722–730, May 2007.

[11] J. Hu, Z. Wang, B. Shen, and H. Gao, "Quantized recursive filtering for a class of nonlinear systems with multiplicative noises and missing measurements," *Int. J. Control*, vol. 86, no. 4, pp. 650–663, 2013.

[12] J. Feng, Z. Wang, and M. Zeng, "Distributed weighted robust Kalman filter fusion for uncertain systems with autocorrelated and cross-correlated noises," *Inf. Fusion*, vol. 14, no. 1, pp. 78–86, Jan. 2013.

[13] J. Feng, Z. Wang, and M. Zeng, "Optimal robust non-fragile Kalman-type recursive filtering with finite-step autocorrelated noises and multiple packet dropouts," *Aerosp. Sci. Technol.*, vol. 15, no. 6, pp. 486–494, Sep. 2011.

[14] S. Cong, J. Layne, and L. Hong, "Iterative robust filtering for ground target tracking," *IET Control Theory Appl.*, vol. 1, no. 1, pp. 372–380, Jan. 2007.

[15] E. Gershon, U. Shaked, and I. Yaesh, "H_∞ control and filtering of discrete-time stochastic systems with multiplicative noise," *Automatica*, vol. 37, no. 3, pp. 409–417, 2001.

[16] F. Yang, Z. Wang, and Y. Hung, "Robust Kalman filtering for discrete time-varying uncertain systems with multiplicative noises," *IEEE Trans. Autom. Control.*, vol. 47, no. 7, pp. 1179–1183, Jul. 2002.

[17] L. El Ghaoui, "State-feedback control of systems with multiplicative noise via linear matrix inequalities," *Syst. Control Lett.*, vol. 24, no. 3, pp. 223–228, Feb. 1995.

[18] M. Sahebsara, T. Chen, and S. L. Shah, "Optimal H_2 filtering with random sensor delay, multiple packet dropout and uncertain observations," *Int. J. Control*, vol. 80, no. 2, pp. 292–301, 2007.

[19] M. Sahebsara, T. Chen, and S. L. Shah, "Optimal H_2 filtering in networked control systems with multiple packet dropout," *IEEE Trans. Autom. Control*, vol. 52, no. 8, pp. 1508–1513, 2007.

[20] W. Deng, J. Yao, and D. Ma, "Time-varying input delay compensation for nonlinear systems with additive disturbance: An output feedback approach," *Int. J. Robust Nonlinear Control*, vol. 28, no. 1, pp. 31–52, Jan. 2018.

[21] W. Deng and J. Yao, "Extended-state-observer-based adaptive control of electro-hydraulic servomechanisms without velocity measurement," *IEEE/ASME Trans. Mechatronics*, to be published, doi: 10.1109/tmech.2019.2959297.

[22] Y. Shi and H. Fang, "Kalman filter-based identification for systems with randomly missing measurements in a network environment," *Int. J. Control*, vol. 83, no. 3, pp. 538–551, Mar. 2010.

[23] F. O. Hounkpevi and E. E. Yaz, "Robust minimum variance linear state estimators for multiple sensors with different failure rates," *Automatica*, vol. 43, no. 7, pp. 1274–1280, Jul. 2007.

[24] J. Linares-Pérez, A. Hermoso-Carazo, R. Caballero-Águila, and J. D. Jiménez-López, "Least-squares linear filtering using observations coming from multiple sensors with one-or two-step random delay," *Signal Process.*, vol. 89, no. 10, pp. 2045–2052, 2009.

[25] X. Lu, H. Zhang, W. Wang, and K.-L. Teo, "Kalman filtering for multiple time-delay systems," *Automatica*, vol. 41, no. 8, pp. 1455–1461, Aug. 2005.

[26] A. Ray, "Output feedback control under randomly varying distributed delays," *J. Guid., Control, Dyn.*, vol. 17, no. 4, pp. 701–711, Jul. 1994.

[27] S. Sun, T. Tian, and L. Honglei, "State estimators for systems with random parameter matrices, stochastic nonlinearities, fading measurements and correlated noises," *Inf. Sci.*, vols. 397–398, pp. 118–136, Aug. 2017.

[28] E. Yaz and A. Ray, "Linear unbiased state estimation for random models with sensor delay," in *Proc. 35th IEEE Conf. Decis. Control*, Dec. 2002, pp. 47–52.

[29] E. Yaz and A. Ray, "Linear unbiased state estimation under randomly varying bounded sensor delay," *Appl. Math. Lett.*, vol. 11, no. 4, pp. 27–32, Jul. 1998.

[30] D. Du, B. Jiang, and P. Shi, "Fault detection for discrete-time switched systems with intermittent measurements," *Int. J. Control*, vol. 85, no. 1, pp. 78–87, 2012.

[31] J. Feng and M. Zeng, "Descriptor recursive estimation for multiple sensors with different delay rates," *Int. J. Control*, vol. 84, no. 3, pp. 584–596, Mar. 2011.

[32] J. Feng, T. Wang, and J. Guo, "Recursive estimation for descriptor systems with multiple packet dropouts and correlated noises," *Aerosp. Sci. Technol.*, vol. 32, no. 1, pp. 200–211, Jan. 2014.

[33] W. Wang and J. Zhou, "Optimal linear filtering design for discrete-time systems with cross-correlated stochastic parameter matrices and noises," *IET Control Theory Appl.*, vol. 11, no. 18, pp. 3353–3362, Dec. 2017.

[34] S. Zhang, Y. Zhao, F. Wu, and J. Zhao, "Robust recursive filtering for uncertain systems with finite-step correlated noises, stochastic nonlinearities and autocorrelated missing measurements," *Aerosp. Sci. Technol.*, vol. 39, pp. 272–280, Dec. 2014.

[35] R. Caballero-Águila, A. Hermoso-Carazo, and J. Linares-Pérez, "Optimal state estimation for networked systems with random parameter matrices, correlated noises and delayed measurements," *Int. J. Gen. Syst.*, vol. 44, no. 2, pp. 142–154, Feb. 2015.

[36] H. Lee, S. Kang, and S. Han, "Real-time optimal state estimation scheme with delayed and periodic measurements," *IEEE Trans. Ind. Electron.*, vol. 65, no. 7, pp. 5970–5978, Jul. 2018.

[37] P. Cui, C. Zhang, H. Zhang, and L. Xie, "Filtering for discrete-time linear noise delay systems," *Int. J. Control Automat. Syst.*, vol. 7, no. 5, pp. 852–857, Oct. 2009.

[38] G. Medrano-Cerda, "Filtering for linear systems involving time delays in the noise process," *IEEE Trans. Autom. Control.*, vol. 28, no. 7, pp. 801–803, Jul. 1983.

JIANXIN FENG received the B.Sc. degree in mathematics from Northeast Normal University, Changchun, China, in 2006, and the M.Sc. and Ph.D. degrees in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2008 and 2012, respectively.

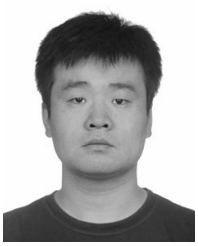
From 2012 to 2016, he was an Assistant Research Fellow with the Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences. Since April 2016, he has been an Associate Professor with the College of Astronautics, Nanjing University of Aeronautics and Astronautics. His main research interests include robust filtering, time delay systems, multisensor data fusion, and line of sight stabilization technology.

RONGNI YANG (Member, IEEE) received the B.S. degree in mathematics from Shandong University, Jinan, China, in 2006, and the M.E. and Ph.D. degrees in control theory and control engineering from the Harbin Institute of Technology, China, in 2008 and 2012, respectively.

From 2009 to 2011, she was a Research Associate with the Faculty of Advanced Technology, University of South Wales, Pontypridd, U.K. From February 2013 to April 2013, she was a Research Associate with the Department of Mechanical Engineering, The University of Hong Kong. From February 2017 to February 2019, she was a Research Fellow with the School of Computing, Engineering and Mathematics, Western Sydney University, Sydney, NSW, Australia. She is currently an Associate Professor with the School of Control Science and Engineering, Shandong University. Her current research interests include networked control systems and multidimensional systems.

Dr. Yang currently serves as an Associate Editor for a number of international journals, including IEEE Access, *Circuits, Systems and Signal Processing*, *Electronics Letters*, and *Neurocomputing*.

HAIYING LIU received the B.Sc., M.Sc., and Ph.D. degrees in precision instruments and machinery from the Nanjing University of Aeronautics and Astronautics, in 2002, 2005, and 2008, respectively. He is currently an Associate Professor and a Master's Tutor with the College of Astronautics, Nanjing University of Aeronautics and Astronautics. His main research interests include navigation, guidance, and control.



BIAO XU received the B.Sc., M.Sc., and Ph.D. degrees in control science and engineering from the Harbin Institute of Technology, Harbin, China, in 2008, 2010, and 2015, respectively. He is currently a Lecturer with the College of Astronautics, Nanjing University of Aeronautics and Astronautics. His main research interests include nonlinear control theory, optimal filtering, aircraft control, and precise guidance.